DOCUMENT RESUME

ED 119 726 IR 003 199

AUTHOR Duchesne, Roderick M.
TITLE Canadian National Bibliographic Data Base Study: The Report.
INSTITUTION National Library of Canada, Ottawa (Ontario). Research and Planning Branch.
PUB DATE Mar 74
NOTE 94p.
AVAILABLE FROM R. M. Duchesne, Research and Planning Branch, National Library of Canada, 395 Wellington Street, Ottawa, Ontario K1A ON4, Canada (Limited quantity).

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage
DESCRIPTORS Bibliographic Citations; Cataloging; *Catalogs; Computers; *Data Bases; Data Processing; Information Processing; Information Retrieval; *Library Automation; *Library Cooperation; Library Networks; Library Planning; Library Technical Processes; National Libraries; *National Programs; On Line Systems
IDENTIFIERS *Canada; Computer Software; Machine Readable Cataloging; MARC; National Library of Canada

ABSTRACT

This study shows strategies which can be used to plan and implement the Canadian National Library's bibliographic data base and the systems which it is to support. The data base would be a subset of a national bibliographic data base which can be brought together in the context of the projected Canadian Library and Information Retrieval Center. Requirements for computer hardware and software are analyzed, and a strategy is outlined for development of the bibliographic data base. (CH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED119726

CANADIAN NATIONAL BIBLIOGRAPHIC DATA BASE STUDY

THE REPORT

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RODERICK M. DUCHESNE

NATIONAL LIBRARY OF CANADA

MARCH 1974

IR003199

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
R.M. Duchesne
National Library of Canada
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

THE REPORT

CONTENTS

	page
Title page	i
Letter of transmittal	ii
Contents	iii
List of Figures	v
List of Tables	vi
List of Annexes	vii
PREFACE	viii
SUMMARY	xi
RECOMMENDATIONS	xv
ABSTRACTS OF ANNEXES TO THE REPORT	xx
I INTRODUCTION	1
II THE CONCEPT OF A NATIONAL BIBLIOGRAPHIC DATA BASE	
A. "Bibliographic data base" : a definition	2
B. The potential national bibliographic data base	2
C. Scope of data base considered in this report	4
III EDP-BASED SYSTEMS AND SERVICES PLANNED BY THE NATIONAL LIBRARY	
A. Description	5
B. Implementation schedule	12

page

IV HARDWARE AND SOFTWARE

A. Hardware 14

B. Software 17

C. Software options 18

D. Factors in the selection of a data base management package 19

E. Data base management packages 21

F. Conclusions of software studies 22

G. Recommendations concerning hardware and software 26

V THE NATIONAL LIBRARY BIBLIOGRAPHIC DATA BASE: BASIC STRATEGY

A. The main current bibliographic record inputs 28

B. Economics of on-line input and storage in a hypothetical CANUC EDP system 30

C. Application of an integrated on-line data base concept 36

D. Order-of-magnitude on-line file size 39

E. Approaches adopted by similar projects elsewhere 41

VI FURTHER ASPECTS

A. Network development 43

B. National Library hardware facility requirements 46

C. Use and treatment of MARC and ISDS records 50

D. Publication of CANUC supplements 51

REFERENCES 57

APPENDICES

1. Economics of on-line input and storage in a hypothetical CANUC EDP system
2. Projected National Library non-Canadiana machine-readable cataloguing 1974-78
3. Union Catalogue of Books: sample layouts used for costing purposes.

LIST OF FIGURES

page

Figure 1 : Diagrammatic representation of the relationship between the National Library bibliographic data base and the national bibliographic data base 4

Figure 2 : Implementation schedule for major NL systems and services 13

Figure 3 : Components of a data base/data communications system 18

Figure 4 : Diagram illustrating the National Library integrated data base concept 38

Figure 5 : Computer configuration illustrating the basic facilities required for adequate support of a National Library on-line integrated bibliographic information system 49

LIST OF TABLES

	page
Table 1 : Main computers presently used by the chief future participants in the Library and Information Retrieval Centre	15
Table 2 : Computers used in library automation in federal government libraries	16
Table 3 : Limitations of existing data base management software packages	24
Table 4 : Data base management software requirements	25
Table 5 : Projected major current bibliographic record inputs to the National Library data base 1974-78	29
Table 6 : Economics of on-line input and storage in a hypothetical CANUC EDP system	32
Table 7 : Illustration of the growth in size of the CANUC and MARC/ISDS portions of an on-line data base given an assumed data base strategy	40
Table 8 : Computer configuration illustrating the basic facilities required for adequate support of a National Library on-line integrated bibliographic information system	48
Table 9 : Theoretical subscription costs for catalogue supplements under alternative assumptions	53

LIST OF ANNEXES

A. Software and file organization sub-study - W.L. Newman and R.G. Smith

B. Model outline of Canadian National Union Catalogue EDP system - R.A. Eldred

C. Canadian on-line cataloguing service model - R.E. Harrison

D. Pattern of use in Canada of Canadian MARC, foreign MARC and ISDS records - S.E. Chapple

E. Projected availability of MARC and ISDS records of prime interest to Canada - R.M. Duchesne and G.V. Ippersiel

F. Canadian National Union Catalogue: projected volumes of accession reports, titles reported and location requests 1974-78 - R.M. Duchesne and G.V. Ippersiel

G. Comparison of MARC monographic formats: Canadian, US, UK and Intermarc - R.M. Duchesne and G.V. Ippersiel

H. France and Sweden: plans for the development of national bibliographic EDP networks - R.M. Duchesne

PREFACE

In July 1973 the National Librarian awarded me a contract with the following terms of reference:

> "Mr. Duchesne will conduct a study of the organization and content of the Canadian National data base of machine-readable bibliographic records and the means by which this may interface with other national and international data bases and machine-readable distribution services. The study will cover the use of various national MARC tape services as input to the Canadian bibliographic data base; the problems of converting other MARC formats to the Canadian MARC format; the dissemination within Canada of records received from foreign tape services to regional centres and libraries; the methods of selecting from such tape services for national distribution; the need for changes and adaptation of foreign records for compatibility with the Canadian format and data base; as well as the organization and accessing of the data base."

Contracts are not easily arranged between parties separated by thousands of miles and I am greatly indebted to the National Librarian, the Associate National Librarian, Miss Hope Clement - and to Mr. A.J. Wells and Mr. R.E. Coward of the British National Bibliography Limited, London, England - for the opportunity to undertake the contract. BNB released me for two-thirds of the study period.

Work on the study commenced in mid-September 1973 and has been essentially a group effort of the Research and Planning Branch of the National Library with copious assistance from within and beyond the Library.

Major direct contributors include:

Mr. W. Newman, Mr. R. Smith - of the National Library Research and Planning Branch	Appendix A
Mr. R. Eldred, Mr. J. Currie, Ms. P. Zuest - all of Digital Methods Limited	Appendix B
Mr. R. Harrison	Appendix C
Ms. S. Chapple, National Library Research and Planning Branch	Appendix D
Mr. G. Ippersiel, National Library Cataloguing Branch; also my full-time assistant on the study	Appendices E and F

National Library staff who have provided invaluable information, comment and/or advice and direction include the National Librarian, the Associate National Librarian, Miss Hope Clement, Miss F. Patterson, Mrs. I. Bradley, Mr. E. Buchinski, Mr. L. Forget, Mrs. J. Giesbrecht, Mr. J. Kelly, Mr. P. Kitchen, Mr. R. Penner, Mr. T. Reid, Mr. A. Schwartz.

The assistance of Dr. J. Brown, Librarian of the National Science Library and his staff is gratefully acknowledged, in particular, information supplied by Mr. P. Wolters, Mr. J. Heilik and Mrs. A. Harvey.

Wider afield I am grateful for information, and in many cases advice, from Mr. B. Stuart Stubbs (Chairman of the Union Catalogue Task Group), Mr. R. Stierwalt (Director, Ontario Universities' Library Co-operative Systems Project), Mr. E. Minet (University of Toronto), Mr. R. Anable (York University), Mrs. H. Avram (Library of Congress), Mr. R. Coward (BNB/British Library), M. M. Boisset (Bureau pour l'automatisation des bibliothèques, Paris), Mme. M. Rosenbaum (Director, International Serials Data Centre), Dr. Ernst

Kohl (Bayerische Staatsbibliothek, München, Germany), Mr. A. Ellis (National Library of Australia), Dr. S. Lindqvist and Dr. M-L. Bachman (Royal Library, Stockholm), Mr. L. Rosquist (Statskonsult, Stockholm), Mrs. I. Elm (Statskontoret, Stockholm), Mrs. A. Dahlstrom and Mr. T. Hedman (Techniska Högskolans Bibliotek, Stockholm).

Not least I wish to record my thanks to Ann Totney my secretary.

Roderick M. Duchesne
29 March 1974

SUMMARY

The present study is essentially a modelling exercise employing available figures and information with order-of-magnitude projections to show the broad strategies which should be followed in the planning and implementation of the National Library's bibliographic data base and the systems which this base is to support over the next five years. This base is a subset of the integrally organized national bibliographic data base which can be brought together in the context of the projected Canadian Library and Information Retrieval Centre.

EDP-based systems and services planned by the National Library (pp.5-13)

In the next two years the National Library will complete the full implementation of its EDP Canadiana systems, and will go on to implement and/or enhance EDP systems in the areas of union catalogues and lists, acquisitions, selective dissemination of information, and - possibly - on-line catalogue support for federal government and other libraries. Most of these EDP systems will centre around the Canadian National Union Catalogue (CANUC) EDP systems, which will cover both serials and non-serials and will provide human-readable union catalogues and lists - as well as on-demand location service similar but faster than that already provided. In systems planning it is already accepted as a principle that new EDP applications should be developed as far as possible as modules of a single overall system.

Hardware and software (pp.14-27)

Hardware availability is a major factor in design of any system with a sizeable on-line data base used many hours a day, and it is of the first importance to National Library system planning to establish the likely future hardware position. Guaranteed access to hardware with suitable operating, communications and data base management software, is a prerequisite for the implementation of an integrated on-line data base serving multiple applications including the CANUC EDP system.

This implies that an extremely high priority should now be accorded to the planning and implementation of the Library and Information Retrieval Centre. The present position of the organizations principally concerned with the Centre is briefly outlined (pp.15-16).

Software requirements for maintenance and use of an integrated on-line data base are analysed. In the context of the Library and Information Retrieval Centre, the National Library will wish to select an operating system and communications monitor likely to be most acceptable and useful to the Centre's participants. Provided that a data base management software package can be found to meet the Library's requirements within acceptable performance, cost and other parameters, there will be very significant economic and elapsed time benefits in making use of such a package. A preliminary survey of National Library data base management requirements and data base management software packages is reported. The conclusion of this survey is that existing software packages do exist which would substantially aid the National Library in the establishment of an integrated bibliographic information system, but that package selection is critically dependent upon the choice of host computer system. More detailed conclusions (pp.22-23) are also drawn, leading to a series of recommendations (pp. 26-27).

The National Library bibliographic data base: basic strategy (pp.28-42)

The key to any integrated EDP system design for the National Library is the CANUC system, since this will have to process a larger volume of records than all the remainder of the National Library's EDP applications put together, when it becomes operational in 1977. The critical feature in the design of the CANUC EDP system is the means by which the input of some 1.7 million accession reports per year is to be achieved. An on-line input system facilitating the use of existing machine-readable data will be much more effici than a batch input system, largely because such a system facilitates the use of existing machine-readable data - and such data will exist for over 88% of repor received, following initial build-up of the EDP file.

The economics of on-line input and storage in a hypothetical CANUC EDP system are examined in some detail, leading to a series of conclusions (pp.33-36). Economic conclusions are dependent on computer hardware being made available at, or near, cost. Principal conclusions are that the National Library should plan to use an integrated on-line data base containing CANUC, National Library acquisitions and in-process, and selected MARC records to support multiple applications. These applications include CANUC accession reporting, CANUC location service, CANUC publication, Canadiana, National Library EDP cataloguing, National Library acquisitions, and catalogue support services for Canadian libraries. Further, all potentially useful records and record data fields should be added to the initial on-line file which should be monitored; following monitoring low-use records and record data fields should be removed from the on-line file as soon as they have been identified.

The outline is drawn of an integrated on-line National Library data base including CANUC, MARC, internal and authority files (pp.36-38). The order-of-magnitude on-line file size is then computed on the basis of a hypothetical data base strategy, showing an on-line file size of under 1,200 million bytes for the first six years of system operation. For purposes of comparison it is noted that the national bibliographic data base of France will be stored on a computer installation with 2,400 million bytes of disk storage. Finally, the basic strategy illustrated is compared with approaches adopted by similar projects elsewhere: national level projects in France and Sweden, and the Ohio College Library Center family of regional projects in the USA. The recommended strategy is very similar to the strategies adopted or planned in these projects. The high cost, size and long-term nature of these projects are noted as well as the fact that problems associated with bilingualism and geographic size will not allow Canadian development to be achieved 'on the cheap'.

Further aspects (pp.43-56)

It is emphasized that the National Library data base and systems strategy will need to be developed within a changing national and international bibliographic EDP network context. Conclusions are quoted (pp.44-45) from a survey of French and Swedish network development plans and from a modelling study designed to

throw light on the economics of alternative on-line catalogue service networks for Canada. Based on these conclusions it is recommended that the National Library, in co-operation with other libraries across Canada, should work out a plan for Canadian national bibliographic EDP network development. It is noted that this recommendation accords closely with the views and recommendations of the Canadian Computer/Communications Task Force. Further, a number of immediate tasks in the network planning area are identified (pp.43-46).

Next, the hardware facility implications of data base and system strategy recommendations are examined. A computer configuration is detailed, illustrating - in the context of the Library and Information Retrieval Centre - the basic facilities required for adequate support of a National Library on-line integrated bibliographic information system (pp.46-49).

The use and treatment of MARC and ISDS in the latter system are then outlined with the conclusion that (selected) ISDS, US MARC, UK MARC and French MARC records should be added to the on-line base. On the basis of current information it appears that foreign MARC tape distribution services within Canada might initially be limited to MARC records of the US, UK and France, and that the preferred format for distribution is the Canadian MARC format. In this connection the conclusions of a substudy are noted: these are that machine translation between the relevant formats should be acceptable for most practical purposes and that translation costs are unlikely to be excessive, although each format translation program will take some man-months to specify, program and bring to operational status (pp.50-51).

Finally, the possibilities are examined for publication of CANUC supplements via a CANUC EDP system. Hard copy publication is not recommended primarily for cost and handling reasons: microfiche or ultrafiche are preferred as forms of publication. Concerning patterns of publication, indexes issued every six months and cumulating over five years are preferred. It is calculated that supplements with a continuously running bibliographic section and indexes of the preferred issue and cumulation pattern could be sold at cost to 300 libraries for an annual subscription of the order of $630 (microfiche) or $750 (ultrafiche) per subscriber. To help provide potential users with a complete picture of the possible economics of this service, suitable microfiche and ultrafiche readers are identified, and their purchase prices as at February 1974 are noted (pp.51-56).

RECOMMENDATIONS

The Library and Information Retrieval Centre

1. Planning for the Library and Information Retrieval Centre should be finalized as a matter of urgency. Speedy determination of the host computer system to be employed by the National Library is a matter of the first importance for National Library EDP system planning purposes. Guaranteed access to hardware with suitable operating, communications and data base management software, is a prerequisite for the implementation of an integrated on-line data base serving multiple applications including the CANUC EDP system. Lack of such access would have severe negative implications for National Library EDP system planning and implementation, would hinder the integration of National Library EDP applications and would critically weaken the economic feasibility and practicality of a CANUC EDP system - the centre-piece of National Library EDP systems to be implemented in the foreseeable future. (p.26)

2. Liaison between the future participants in the Centre should be strengthened to assist planning, to ensure maximum compatibility and to help the sharing of information and experience. Liaison concerning software development is of particular importance. The various participants in the Centre may commit themselves to different operating systems, communication monitor and data base management software, hindering future operation on a common facility and hindering future integration of files and systems. While a single set of software may not meet all the needs of each participant, liaison should assist agreement on at least common operating system and communication monitor software and on preferred programming languages. (p.26)

The Canadian national bibliographic EDP network

3. The National Library, in co-operation with other libraries across Canada, should work out a plan for Canadian national bibliographic EDP network development. This plan should be phased and should show for each phase the extent of the network, including:
   - details of participating organizations
   - location and nature of hardware employed
   - location and nature of files accessible to the network
   - communications facilities
   - services and facilities offered by each participating organization

   This recommendation accords closely with the views of the Canadian Computer/Communications Task Force, exemplified in the following recommendations of the report of the Task Force entitled Branching Out [27]:

   1. "Computer/communications (i.e. computer services by remote-access through communications facilities) should be recognized by governments as a key area of industrial and social activity, and steps should be taken towards strengthening ... co-ordination of its development to the benefit of Canadian society."

   2. "In the formulation of national computer/communications policy a unified approach throughout Canada should be stressed as a key factor requiring close co-ordination between federal and provincial actions."

   3. "In the area of federal responsibilities a Focal Point should be established within the government for co-ordination in the development, formulation and continuing evaluation of national policy in all matters pertaining to the field of computer/communications."

Immediate tasks in the area of bibliographic EDP network planning include:

- the strengthening of organizational arrangements and the provision of additional resources to support joint national/provincial planning concerning library and information service development.

- Canadian National Union Catalogue system planning, for instance the manner in which libraries will report when the new CANUC EDP system is operational. While many libraries may wish to continue reporting exactly as they do now, some will wish to report in machine-readable form. The mode and timescale of this machine-readable reporting needs to be worked out in detail.

- Planning of catalogue support services, both off-line and on-line. Joint national/provincial planning is required since some services will be provided by the National Library, some will be provided by inter-provincial and intra-provincial projects (for example Ontario Universities' Library Co-operative System) and some may be provided by foreign or international projects, for example Ohio College Library Center.

- Active participation in international bibliographic EDP network planning and development, for instance MARC, ISDS and UNISIST network planning and development. (pp.45-46)

National Library data base strategy

4. Given adequate and economic host computer hardware availability the National Library should plan to use an integrated on-line data base containing CANUC, National Library acquisitions and in-process files, and selected MARC records, in which the data of each bibliographic item is held only once. This base can serve many different applications including CANUC accession reporting, CANUC location service, CANUC publication, Canadiana, National Library EDP cataloguing, National Library acquisitions, and

catalogue support services for Canadian libraries. All potentially useful records and record data fields should be added to the initial on-line file, and monitoring should be undertaken to identify low use records and data fields. Low-use records and record data fields should be removed from the on-line file as soon as they have been identified by the monitoring process. (p.35)

This strategy accords with that adopted or planned by other major data base projects, notably the national data base projects of France and Sweden and the Ohio College Library Center family of projects in the USA. As with these projects elsewhere, full development of the Canadian national bibliographic data base and the services supported by this base will be a major enterprise involving considerable expenditures over the greater part of a decade. Special Canadian requirements, such as those associated with bilingualism and network development in a country of Canada's geographical size, will add to the already complex problems related to the development of a large national bibliographic data base and the services to be provided from this base. The gains from undertaking this type of development may be inferred from the way in which the OCLC system is being replicated in different regions of the USA on a self-supporting basis and from the national plans of France and Sweden, as well as the more direct National Library, CANUC and other benefits which may be realized in Canada. (pp.41-42)

National Library software development

5. The National Library should specify in more detail the data base structure required for its purposes and should carry the study of data base management software packages to the benchmark testing stage. Data base management software, such as CAN/OLE, presently used or being developed by other Library and Information Retrieval Centre participants should be studied before final selection of software for benchmark testing. (pp.26-27)

6. Following testing and selection of data base management software, similar detailed analysis including benchmark tests should be performed in respect of communications monitor packages. (p.27)

7. Following these stages - and given adequate and economic host computer system availability - the National Library should proceed to implement an integrated data base employing defined structures and chosen data base management and communications monitor software. (p.27)

ABSTRACTS OF ANNEXES TO THE REPORT

A. Software and file organization (68p)*

The purpose of this sub-study is to determine whether any existing data base management ('DBMS') and communications monitor software packages have facilities which satisfy the requirements of the National Library. The general characteristics of these types of software are described including desirable features. An outline of the National Library data base is then presented including its structure, levels of bibliographic record and desirable linkages. Linkages include 'vertical' (e.g. collection, unit, analytic), 'horizontal' (e.g. copy and part), and 'parallel' (e.g. supplements, continuations, reprints, translations). Access methods are examined including compression code, word, and normalized entry access, and access points needed for different files and levels of bibliographic item in the National Library data base.

National Library requirements of DBMS are then considered in more detail and the results of a preliminary literature survey of 13 packages are reported. Two DBMS and associated communications monitor packages which (on paper) satisfy National Library requirements are reported in more detail. Finally, a series of conclusions and recommendations are drawn: these are included in Chapter IV of the main report.

B. Model outline of Canadian National Union Catalogue EDP system (80p)

Following presentation of background facts, statistics and assumptions, assumed CANUC EDP record contents and layout and access methods are presented. Average record size and file size growth are computed employing stated

---

* Page totals shown for each annex in this section include front matter, text, and appendices.

assumptions. Data base software assumptions are stated and terminal numbers and unit costs for accession report input and for location searching are computed. An illustrative computer hardware installation is presented and it is noted that a number of commercial bureaux have large enough installations to support - at least temporarily, say in early testing stages - a CANUC EDP system (bureau economics are not examined). Conclusions concerning the CANUC EDP system are then presented.

The remainder of the annex deals with publication of CANUC catalogues employing the EDP system. Assumed volumes, layouts and alternative publication patterns are outlined, followed by a description of publication media and Computer Output Microfilm ('COM') systems. Suitable microfiche and ultrafiche readers are identified, and subscription costs for assumed EDP system-produced catalogues are computed. Finally, a possible Voice Answerback telephone location service is outlined.

C. Canadian on-line cataloguing service model (75p)

The purpose of this study is to throw light on the economics of national versus regional provision of on-line catalogue support services. Costings are based on Ohio College Library Center 1971/72 costs with networking costs provided by the Computer Communications Group of the Trans-Canada Telephone system and Xerox of Canada Limited hardware costing of an OCLC configuration. National versus regional costs are computed for two networks chosen to be approximately the same size as the OCLC network at the time of study work, and approximately twice this size. The main conclusions are reported in Chapter VI of the main report.

D. Pattern of use in Canada of Canadian MARC, foreign MARC and ISDS records (28p)

The purpose of this annex is to summarize known requirements of Canadian libraries for MARC and ISDS records. The present extent of the Canadian MARC record distribution service and its future prospects are described. Known use and intended applications for Canadian MARC tapes are summarized, including the responses to a poll of Canadian MARC pilot project recipients concerning frequency and types of service required. Canadian MARC format development plans and service plans are noted. Canadian MARC pilot project participants' format preferences and choice of tapes and distribution frequency are outlined.

Tables show Canadian MARC pilot project participants' known applications, the Canadian MARC format development schedule, participants' hardware, and an outline of the topics suggested for inclusion in the participants' reports at the end of the pilot project.

E. Projected availability of MARC and ISDS records of prime interest to Canada (30p)

Availability of these records is relevant to the overall study because they can be used both for catalogue support purposes and to reduce data input costs; if the records are added to the on-line data base their number and storage requirements are of importance in data base design.

Based on information supplied by the national and international agencies concerned, annual production of current records is projected for all ISDS records and MARC records from the following countries: Australia, Canada, France, Germany, United Kingdom, and the United States. Projections are shown for each service both for the five calendar years commencing 1st January 1974, and the five National Library fiscal years commencing 1st April 1974. Retrospective records projected to be available in the period 1974-78 are detailed separately.

Based on stated assumptions the number of current MARC and ISDS titles for the same group of services is calculated, giving one total figure for ISDS and MARC services of France, UK and the USA, and one grand total figure for the whole group. These two figures of current titles are projected for the thirteen year period to 31st December 1986; the period expected to cover the first ten years of life of the CANUC EDP system.

F. Canadian National Union Catalogue: projected volumes of accession reports, titles reported and location requests 1974-78 (14p)

These projections are of fundamental importance to National Library data base and CANUC EDP system design. National Library union catalogues are briefly described. CANUC accession report statistics are projected to 31st March 1979 based on historical statistics commencing 1st April 1960. Statistics are divided into the following groups: "non-serials", "serials other than newspapers" and "total" - the latter being the combination of the two former classes. Newspaper accession reports are not included because the volume of current reports is very small. CANUC titles reported are projected in the same groups over the same timescales. Location requests are projected over the same timescales employing the same groups, with the addition of newspaper location requests which are shown separately and added into total figures. Lastly, a summary graph is presented of historical and projected statistics for all material for the ten National Library fiscal years to 1978/79.

G. Comparison of MARC monographic formats: Canadian, US, UK and Intermarc (54p)

This annex presents a comparative table of content designators of the monographic communication formats indicated in its title, and draws conclusions based on this table and the operational experience of the British National Bibliography Limited, London, England.

H. France and Sweden: plans for the development of national bibliographic EDP networks (29p)

This annex records information obtained from visits to the countries and from the literature, and draws overall conclusions which are quoted in Chapter VI of the main report. The section concerned with France sketches the organization of libraries in France, and describes both the Bureau pour l'automatisation des bibliothèques and the remaining three years of the sixth national five-year plan for library automation. A bibliography of relevant literature is provided. The section concerned with Sweden is similarly structured. Draft documents were circulated to persons in France and Sweden and the present annex incorporates changes suggested by these persons.

I. INTRODUCTION

Definitive planning of the organization, structure, and modes of access to be employed in a machine-readable data base requires precise knowledge of the following:

1. Services and outputs to be provided from the data base, including details of transaction volumes and response times over the life of the base.

2. Inputs to the data base including not only crude volumes, but also the degree of duplication of records and of fields within records.

3. Hardware to be employed.

It commonly takes about 3 years to bring a large EDP application to operational status and the system installed commonly has a life of about 7 years. From this it is seen that data base planning has to look some 10 years into the future. Clearly there is a high degree of uncertainty in such a long look ahead, compounded by the fact that the major design-determining factors noted above represent at this time questions without answers adequately defined for definitive data base design purposes.

In these circumstances the present study is essentially a modelling exercise employing available figures and information with order-of-magnitude projections to show the broad data base and system strategies which should be followed. This point is emphasized: the purpose of the report is to help determine strategy. Matters of detail are included insofar as they are pertinent to strategy: many matters of detail are not treated because they are not critical in system design.

II. THE CONCEPT OF A NATIONAL BIBLIOGRAPHIC DATA BASE

A. "BIBLIOGRAPHIC DATA BASE": A DEFINITION

"Bibliographic" is used in this report in the sense of "pertaining to library materials". These materials include monographs, serials, theses, maps, music and audio-visual materials. This list is not exhaustive and in the medium or longer term it may be advantageous to include documentation, library and archive material EDP records in a single data base system.

A "data base" for the purposes of this report is "a sizeable collection of computer-readable data organized for use". The base may comprise a number of files held on a variety of storage media. For example, some organizations with disk data bases and the need to publish large volumes of records maintain a separate formatted output file to save on publication processing costs. In this case, multiple copies of a record exist for a single item: a master disk record (not necessarily held as one integral record on disk), at least one copy of this record on tape for regeneration purposes, at least one formatted output record, and at least one copy of the formatted output record for regeneration purposes.

B. THE POTENTIAL NATIONAL BIBLIOGRAPHIC DATA BASE

The Treasury Board EDP master plan[1] projects the creation of a Library and Information Retrieval Centre under the custodianship of the National Librarian. Principal departments recommended to participate in the Centre are: National Library; National Research Council (National Science Library); Public Archives; National Museums; Information Canada; Defence Research Board; Agriculture.

The National Library and National Science Library are envisaged as "the nucleus of a service-wide application centre for information retrieval"[2] taking into account the fact that these two libraries have "given rather careful study to the establishment of a joint computing facility which would meet, eventually, the needs of all government libraries in the information retrieval area"[3].

At the present time the departments which will participate in the Centre maintain separate files, and it is not very meaningful to refer to the totality of their machine readable files as a single data base. Insofar as their existing files of machine-readable bibliographic records are large enough to merit the term "bibliographic data base", they exist as separate data bases. However, when the files are processed on a single installation it will be possible to integrate them into a single organized system. When this has been done it will be meaningful - with one qualification - to describe these machine-readable files relating to library materials as "the national bibliographic data base".

The qualification relates to the medium or longer term in which an on-line national bibliographic network is envisaged. At that time it will be possible to integrate machine-readable files at remote locations into a single organized system. The national bibliographic data base will then comprise all the accessible bibliographic files at the nodes in the national network.

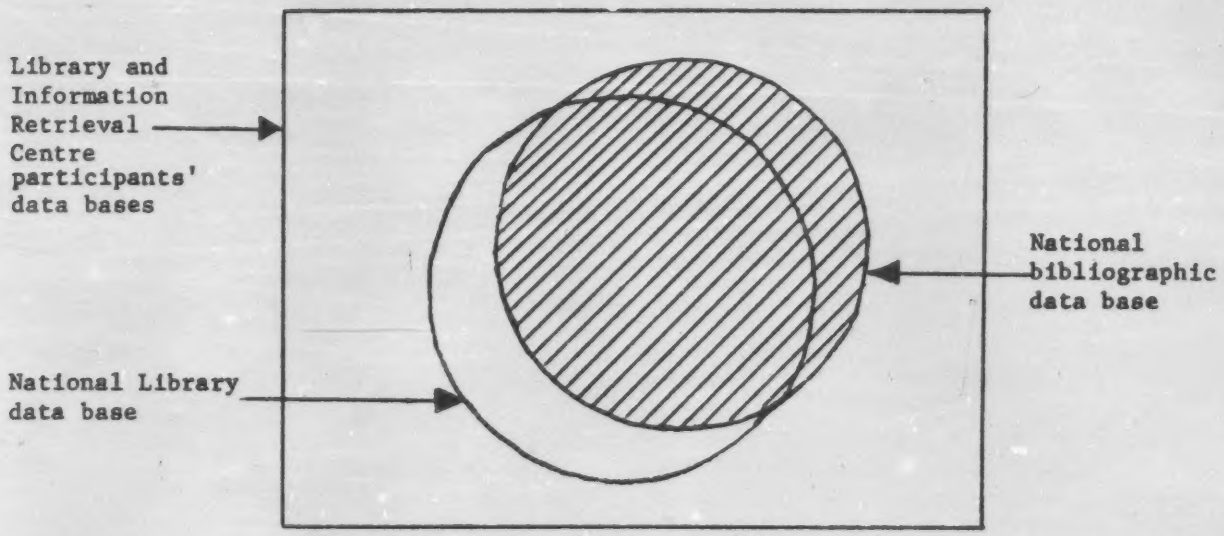
From these observations it is clear that "the national bibliographic data base" does not at present exist in an integrally organized form. Further, when the integrally organized base comes into existence, its initial scope will be much more restricted than its likely eventual scope. The scope of the data base considered in this report is indicated in the next section.

C. SCOPE OF DATA BASE CONSIDERED IN THIS REPORT

This study is concerned with the development of the National Library bibliographic data base over the next 5 to 10 years. This may be regarded as one of the necessary steps in data base planning for the Library and Information Retrieval Centre: a realistic approach is to consider the plans of each of the proposed participants in the Centre and use the resulting information to consider to what extent these plans can be integrated.

The relationship of the National Library data base to the national bibliographic data base and its associated non-bibliographic files can be depicted diagrammatically. Figure 1 below depicts the situation in which the Library and Information Retrieval Centre has been set up and holds the national bibliographic data base.

Figure 1 : Diagrammatic representation of the relationship between the National Library bibliographic data base and the national bibliographic data base


Library and Information Retrieval Centre participants' data bases
National bibliographic data base
National Library data base

Legend

- bibliographic files relating to library, documentation and archive material
- non-bibliographic files, for example mailing lists, patron files, vendor files

III. EDP-BASED SYSTEMS AND SERVICES PLANNED BY THE NATIONAL LIBRARY

A. DESCRIPTION

The function of the National Library data base is to support the EDP operations of the Library. Existing, projected and likely EDP systems and projects are noted below, grouped into system areas. This information is drawn from plans[4] endorsed by the Library's Committee on Automation and by the Associate National Librarian in January/February 1974. These plans are, however, subject to Treasury Board approval and to the availability of the hardware, staff and other resources needed for their realization. The general philosophy behind these plans is encapsulated in the first recommendation of the Report of the Systems Development Project[5]:

> "An integrated bibliographic information system for the National Library be developed comprising all the operations of acquisition, cataloguing, Canadiana, serials control, the Union Catalogue and union lists which permit handling by electronic data processing. The integrated system will operate in a combination of batch and real-time modes."[6]

1. Canadiana

Plans in relation to Canadiana have been announced in National Library News November-December 1972[7] and January 1974[8]. The plans call for:

(a) Continuation of the EDP-based cataloguing and publication system for monographs; extension of this system to cover firstly serials and government documents and secondly, audio-visual materials. Products include the national bibliography Canadiana and a proof service to libraries for assistance in their technical processing.

(b) Use of the EDP-based Canadiana system to produce National Library catalogue cards for the Library's Canadiana (and non-Canadiana) intake.

(c) For Canadian serials: production of machine-readable records conforming to ISDS standards for submission to the International Serials Data Centre, Paris.

(d) For government documents: provision of output from the National Library data base to a government documents indexing subsystem - almost certainly that developed at the University of Guelph. The possibility of producing the Information Canada Daily Checklist from Canadiana records is presently being studied by a task group.

(e) Production and distribution of Canadian MARC records.

(f) Retrospective conversion of all records for Canadian books published before 1949. First priority is given to all records for the bibliography 1867-1900, in order to publish a bibliography covering these years.

(g) Eventually, following investigation, the Preserved Context Index System (PRECIS) may be implemented to provide a subject index to Canadiana and to provide additional subject access to MARC records.

2. National Library cataloguing

(a) National Library Canadiana cataloguing has been dealt with above under the heading of Canadiana. NL non-Canadiana cataloguing is to be progressively transferred from manual to EDP procedures within the next three years. The EDP procedures are those presently employed by the EDP-based Canadiana system.

(b) Participation in Ontario Universities' Library Co-operative System (OULCS) has been arranged for experimental purposes. It is probable that the system will be used to catalogue a quantity of foreign books.

3. SDI

SDI services are currently provided in co-operation with the National Science Library against the following magnetic tape services:

US MARC	-	all subjects
ERICTAPES	-	education
SSCI	-	Social Sciences Citation Index

It is proposed to search additional tapes:

Canadian MARC)	1974
Psychological Abstracts)	
Historical Abstracts	1975

Further tapes are under consideration: PREDICAST (marketing information) Sociological Abstracts; UK MARC tapes; French MARC tapes; PASCAL (Centre National de la recherche scientifique - Bulletin signalétique).

Concerning SDI, the following recommendations of the Federal Government Library Survey are relevant:

(a) "Priority be given to the development of CAN/SDI services in the social sciences and humanities to parallel services provided in science and technology"[9].

(b) "CAN/SDI be expanded to include all available tape services, social sciences and humanities as well as scientific and technical, and that the service and its training programs be made available to all federal libraries through the Library and Information Retrieval Centre"[10].

4. Canadian National Union Catalogue (CANUC)

Policy with respect to this catalogue and a "New Automated Union Catalogue of Books" was announced in National Library News, January 1974[8]. The National Library is to start a new EDP-based union catalogue system and will close the existing Union Catalogue of Books ('UCB') as soon as the new system is operational. Following editing, the UCB will be published, probably in microform. The new system will provide the facility of publishing the catalogue at intervals. Access to the new EDP catalogue within the National Library will probably be on-line.

At a later date the CANUC EDP system may be developed to provide statistical analyses of the items reported to CANUC; in particular, the resources survey of university library collections in Canada could be kept up to date in this way.

5. Union Lists

(a) Serials

As noted in National Library News, January 1974, "it is not possible to make any decisions in this area at present because a great deal of work is in progress"[11]. If the interim recommendations of the Canadian Union Catalogue Task Group are accepted, the National Library will "proceed quickly with the development of a union list of periodicals in the humanities and social sciences as a first priority, and of newspapers as a second priority"[12]. The latter recommendation accords closely with the recommendation of the Federal Government Libraries Survey that "The National Library give priority to the compilation of the computerized union list of serials in the social sciences and humanities"[13].

If the recommendations of a recent report[14] are accepted, the Canadian Union Serials List (CUSL) will be "produced in book form at least every two or three years in toto, plus an accumulated supplement at least every six months"[15]. Also the National Library and National Science Library will "formally organize to jointly produce and maintain a single Canadian Union File of Serials (CUFS) covering all subjects"[15] and these libraries will "participate in the proposed CONSER* Project"[16]. Work in co-operation with the National Science Library is proceeding in the serials area.

(b) Other material

Following the production of the Canadian Union Serials List, further union lists are envisaged, but these have not yet been accorded the status of definite EDP-based projects in the current development schedule[4]:

Union catalogue of audio-visual materials
Union catalogue of materials for the blind
Union list of newspapers

6. Integrated system support

A machine-readable authority file for government headings will form part of the Canadiana system for serials and government documents to be operational in November 1974. It is planned to extend this file to all types of heading including personal, corporate, conference, and series headings. The authority file will include headings in Canadiana, National Library cataloguing - and in course of time - CANUC, and National Library acquisition and serials control files. If a recommendation of the Federal Government Library Survey is accepted "the authority file for personal and corporate authors" will be published and frequently updated[17].

---

* Co-operative project for the CONversion of SERials

Creation of a machine-readable authority file is tentatively scheduled for 1975/76. Also pertinent to integrated system support is a proposal to investigate data base management and communications monitor software.

7. MARC and catalogue support services

The National Librarian has announced (January 1974)[18]:

"Next month, the National Library will begin to produce Canadian MARC tapes, and the library intends to exchange these with other countries in return for their tapes, as well as to purchase other tapes from various countries and will be able to provide a machine-readable service in a batch mode, providing selected records on magnetic tape at cost to interested libraries and bibliographic centres. Libraries could subscribe to all the records from a country, to records selected by categories, or to specific records selected by numbers such as the LC Card Number or the International Standard Book or Serial Number. This will form an off-line cataloguing support service which will provide subscribers with machine-readable records for use in their own systems.

"First priority is to be given by the National Library to implementing the new automated system for the union catalogue and to developing the bibliographic services using MARC tapes. The second priority is to develop for federal government libraries a cataloguing service which may be in the form of a centralized, cooperative or shared cataloguing service with the possibility of the system being extended to other libraries in the country which wish to participate. Next year, the National Library will investigate the best means of providing cataloguing support to the federal government libraries."

8. Acquisitions

(a) All materials

In due course the National Library will develop an EDP-based subsystem supporting selection, ordering and claiming for all materials. The subsystem will interface with cataloguing and catalogue support subsystems: bibliographic information captured in the acquisitions process will be available for use in the cataloguing process. Conversely, machine-readable data from CANUC and MARC files will be used in the acquisitions process. A centralized acquisitions system for all Federal Government libraries is a possible later development.

(b) Serials control

Serials control automation will be concerned with the creation and maintenance of a serials control file against which incoming issues may be checked and which will produce various lists of serials holdings as well as immediate information on the receipt of a serial issue by the National Library. Claiming, binding, renewal and routing functions may be included in this system, which may eventually be extended to Federal Government libraries.

9. Multilingual Biblioservice

The automation requirements of the Biblioservice are not yet fully defined. When operational, the Biblioservice will rotate book collections in other than the official languages among public and regional libraries; these will, in turn, make them available to interested people in the areas they serve. Following more detailed investigation, an EDP system may be devised to support this operation.

10. Directories and mailing lists

It is planned to study the feasibility of employing EDP means to create, maintain and publish various Canadian library directories presently prepared or being considered by the Library Documentation Centre. The various mailing lists required by the National Library might be produced from the file used to produce the directories.

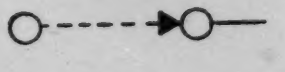
11. Canadian Book Exchange Centre

It is planned to investigate the feasibility of an EDP system to create, update and distribute lists of surplus materials.

B. IMPLEMENTATION SCHEDULE

Figure 2 below shows the timescale of the implementation of the systems and projects outlined above. It has the following legend:

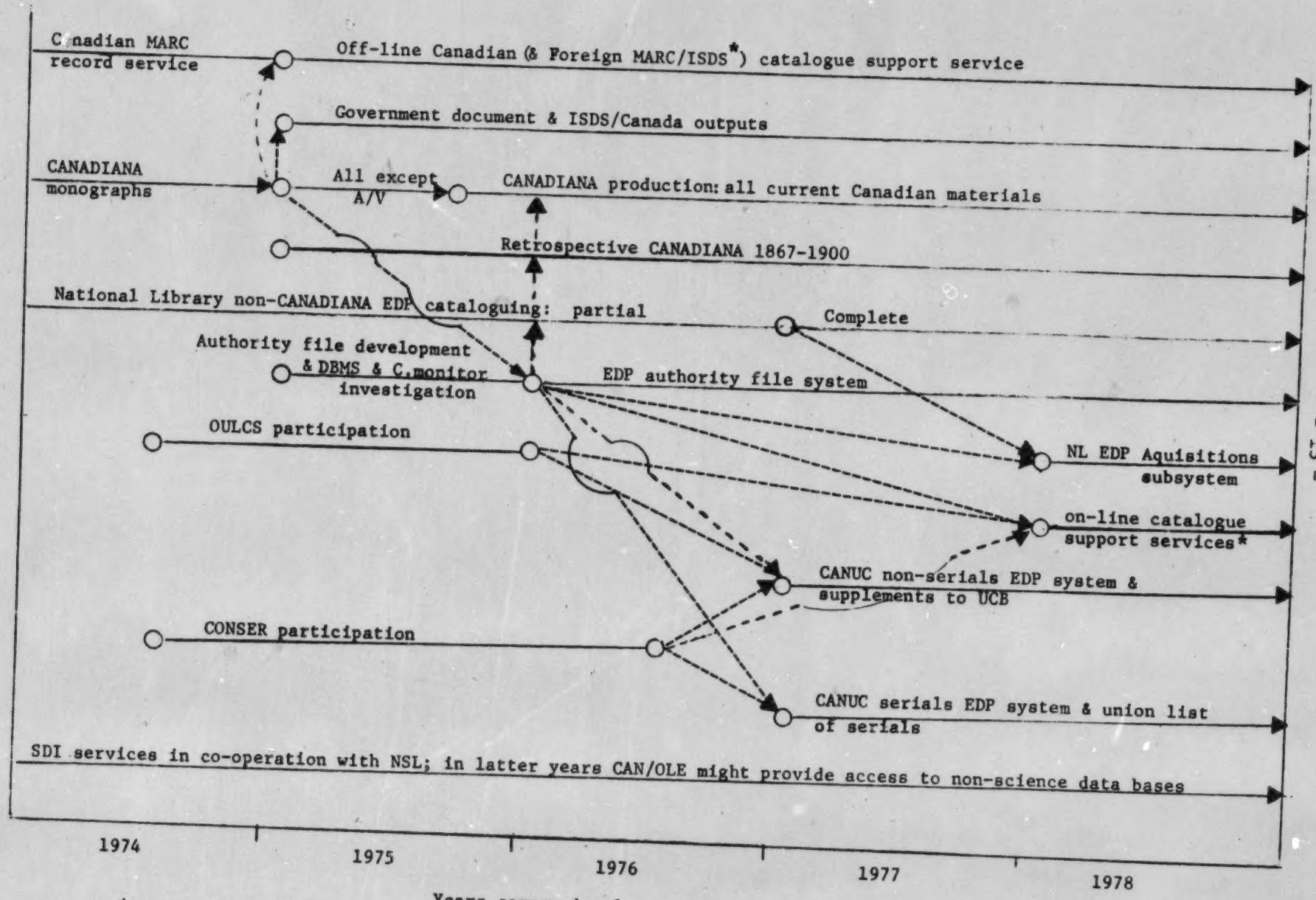
Commencement of new or enhanced system or service



Commencement as noted above is dependent on or closely linked with some other system or service

Service or system continues beyond timescale of chart

Figure 2 : Implementation schedule for major NL systems and services


Canadian MARC record service
Off-line Canadian (& Foreign MARC/ISDS*) catalogue support service
Government document & ISDS/Canada outputs
CANADIANA monographs
All except A/V
CANADIANA production: all current Canadian materials
Retrospective CANADIANA 1867-1900
National Library non-CANADIANA EDP cataloguing: partial
Complete
Authority file development & DBMS & C.monitor investigation
EDP authority file system
OULCS participation
NL EDP Aquisitions subsystem
on-line catalogue support services*
CANUC non-serials EDP system & supplements to UCB
CONSER participation
CANUC serials EDP system & union list of serials
SDI services in co-operation with NSL; in latter years CAN/OLE might provide access to non-science data bases
1974
1975
1976
1977
1978
Years commencing 1st April

* Not yet definite projects

IV. HARDWARE AND SOFTWARE

A. HARDWARE

Although the National Library is committed[3] to participate in the Library and Information Retrieval Centre, the plans have not reached the stage at which the future hardware of the Centre, and the timescale of the availability of this hardware, are determined.

Hardware availability is a major factor in the design of any system with a sizeable on-line data base used many hours a day. If the system has tc be operated for a significant proportion of its life in a bureau environment at bureau prices, this is likely to have a considerable impact on systems, procedural and data base strategies. For example, it may not be economic to maintain a very large data base on-line for all normal working day hours in a bureau environment. In this case it may be necessary to carry out more processing off-line.

It is therefore of the first importance to National Library system planning to establish the likely future hardware position. This implies that an extremely high priority should now be accorded to the planning and implementation of the Library and Information Retrieval Centre. Important steps will be to:

1. Set up a study with appropriate terms of reference, and controlling and liaison arrangements
2. Study the requirements of the participants in the Centre
3. Report with recommendations
4. Consider the report and recommendations
5. Decide policy and implementation plan
6. Proceed with implementation

These steps will not be accomplished overnight, and the future participants in the Centre will not be able to hold back their system development to await decisions concerning the Centre. In this situation there will be considerable advantages in strengthening liaison and co-ordination between the participants to ensure that their developments are, and remain, as compatible as possible. The software aspects of this argument are developed in more detail in later sections concerned with software. Present hardware compatibility is generally indicated in the two tables below.

Table 1 : Main computers presently used by the chief future participants in the Library and Information Retrieval Centre

Chief participants as listed in EDP Master Plan	Computers[a]	Major Operating Systems
National Library	IBM 360/40 IBM 360/50	OS/MVT
National Research Council (National Science Library)	IBM 360/67	TSS
Public Archives	IBM 360/50 IBM 360/40	DOS, transferring to OS
National Museums	IBM 370/135 IBM 360/85 IBM 360/50 PDP 11/05	OS/VS OS
Information Canada	IBM 360/65 RCA Spectra 7045	OS
Defence Research Board	Varian 620L IBM 370/145 IBM 370/165	OS
Agriculture	IBM 360/50 IBM 360/40 IBM 370/165 Univac 1108	DOS, transferring to OS DOS and OS Exec 8

Source: Canadian Library Directory 1974 and direct contact of the organizations.

Notes: (a) Most of the computers shown belong to service bureaux.

Federal Government libraries would also be prospective users of the Centre. Table 2 below shows the present position with regard to these libraries.

Table 2 : Computers used in library automation in federal government libraries

Computer	No. of computers	No. of systems
IBM 1401	1	1
360/30	1	1
360/40	2	6
360/50	2	2
360/65	3	8
360/67	1	5
CDC 6600	2	1
Univac	1	1
Ferranti/Packard	1	1
CDC 3100	1	1
Totals	15	31

Source: Federal Government Library Survey; report of the Services and Systems Team. Ottawa, National Library, 1973, p.267

B. SOFTWARE

As noted in Annex A, four categories of software are required to maintain and use an integrated on-line data base, including terminal access to this base:

1. User application programs

These provide for the detailed processing of records. Examples of this processing are validation of input records and formatting of records for printing. Utility programs for sorting, merging and dumping are included in this category.

2. Data base management

This software maintains the files and their indexes, and provides user application program and communications monitor access to these.

3. Communications monitor

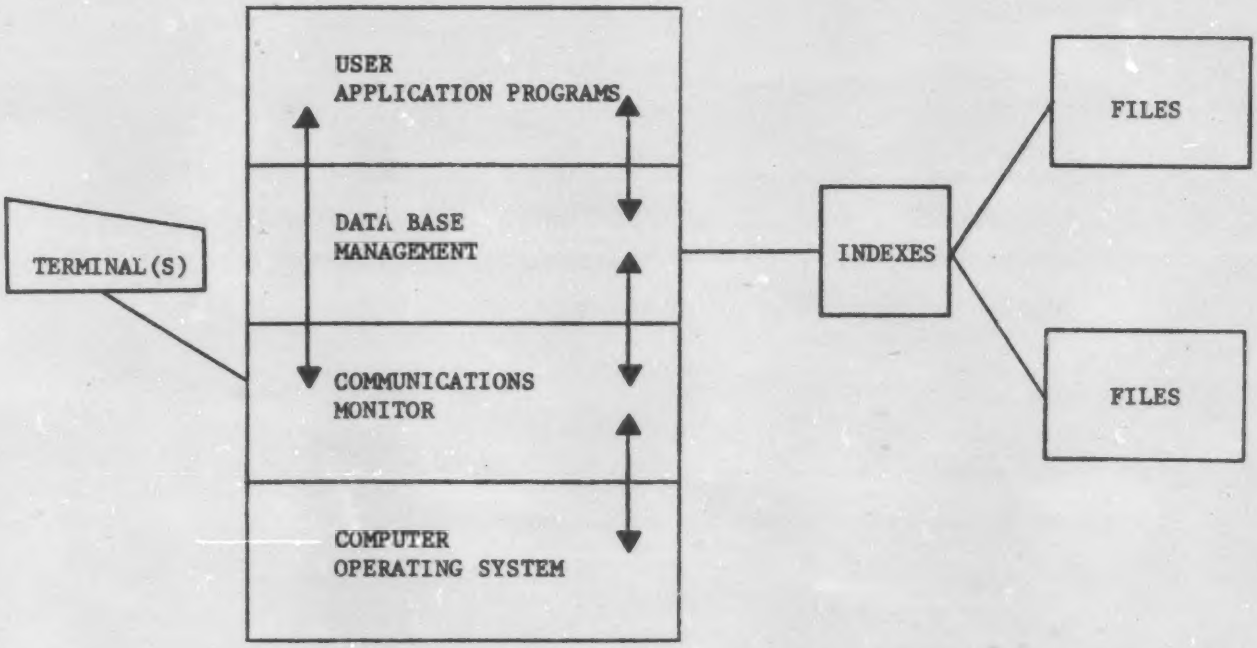
This software "acts as a three-way interface between messages from teleprocessing terminals, the application programs which determine how the messages are to be processed, and the operating system. It controls the queuing, editing, and dispatching of the messages, the centralized input/output operations involved, and the rolling in and out of programs"[19].

4. Computer operating system

This provides overall control of operations including allocation of core storage and peripheral devices to programs.

Figure 3 below illustrates the interrelationships of these categories of software in a data base/data communications system.

Figure 3 : Components of a data base/data communications system



C. SOFTWARE OPTIONS

In the context of the Library and Information Retrieval Centre, the National Library will wish to select an operating system and a communications monitor likely to be most acceptable and useful to the Centre's participants. It is hardly conceivable that this choice could involve modification of a manufacturer's operating system or writing communications monitor software.

The position with regard to data base management software is also clear-cut, provided that an existing package can be found which meets the Library's requirements within acceptable performance, cost and other parameters. The economic and elapsed time benefits of making use of an existing data base management system (DBMS) are indicated in the following quotation from Annex A:

> "The University of Chicago[20] estimated that the development of a sophisticated access method (the basis of a DBMS) would require 86 man months, and would cost $151,000. Note that an access method is not a full DBMS, and that a satisfactory product was not guaranteed. MRI Systems Corporation[21] quotes studies made by a number of management consulting firms which price in-house development of a sophisticated DBMS, such as System 2000, at $1.5 million to $3 million. Purchase of a proprietary DBMS package ... is obviously less expensive than in-house development."

Factors in the selection of data base management software packages are discussed below.

D. FACTORS IN THE SELECTION OF A DATA BASE MANAGEMENT PACKAGE

The list of factors below summarizes the salient factors applicable in the case of the National Library; it is drawn from a fuller analysis contained in Annex A:

1. Hardware and operating system

The package must operate on the hardware available under an acceptable operating system.

2. Facilities

The package must provide the basic facilities required. In the present case it must be capable of maintaining files of sizes of a few hundred up to 10 million variable length records with a varying number of variable length fields. Access via multiple record keys must be possible (e.g. LC number, ISBN, author and title compression codes) with provision of pre-search statistics. It must be able to store portions of the data base on sequential storage devices, such as magnetic tape. Facilities must be available for linking records and parts of records as required. Linking and association requirements are analysed in some detail in Annex A; they are rather more complex than in the average data base. In addition, facilities for regenerating data in the event of hardware or system failure are mandatory, and there should also be facilities to prevent access by unauthorized persons to classified data.

3. Dependability and transferability

The package must be operational and proven on as many installations as possible. It should be maintained and updated by a reputable and reliable software development organization. Updating should also ensure that package users can make use of new versions of the same operating system and new models in the same hardware range.

4. Flexibility

The package should allow reorganization of the data base, creation and deletion of files, and amendment of record structures without the necessity for package reprogramming.

5. Performance and cost

The package should not be unduly costly to rent (or buy) and implement. Once implemented it should operate within acceptable core storage limits and should maintain indexes and files within acceptable storage limits and costs. In operating it should provide timely response to user requests without excessive "housekeeping" costs. Housekeeping in this context refers to operations such as file reorganization, index maintenance and updating, generation of links between records and parts of records. There is in practice often a trade-off between response time and housekeeping cost: faster response time may be achieved at the expense of increasing housekeeping overhead. Alternatively, housekeeping costs may be lowered at the expense of accepting a slower response time.

E. DATA BASE MANAGEMENT PACKAGES

The list of factors just presented shows that package selection is not a simple task: it involves weighing many factors. Only limited conclusions can be drawn on the basis of published literature and specifications. All that can be decided on the basis of these is whether there appear to be suitable and acceptable packages and which packages should be tested out in practice. Empirical test is mandatory in order to establish factors such as response time and storage requirements in a practical context.

The two tables below are drawn from Annex A; they show the range of packages given preliminary study, and summarize the findings. Within the limited time available it was not possible to study every package in existence; omission of a given package from the table does not mean that this package has been studied and definitely rejected. The objective was to carry out a preliminary survey of the main available and eligible packages.

CAN/OLE could not be studied in detail at the time of this preliminary work since it was in the process of development at the National Science Library, and documentation concerning its use and capabilities was not available; it should definitely be included in the next round of studies.

ISIS should also be considered, especially if it is selected by National Museums; ISIS is presently used by the Swedish LIBRIS project and was developed at the International Labour Office, Geneva. It is presently supported by the International Development Research Centre in Ottawa and is being developed to enable it to handle larger data bases.

STAIRS is an information retrieval package rather than a comprehensive data base management system. For instance, it does not include an on-line updating facility. Like LIBRIS, it would also need to be considered in more detail, especially if selected by National Museums.

F. CONCLUSIONS OF SOFTWARE STUDIES

The conclusions of work done in the software area are as follows:

1. Data base management systems and communications monitors do exist which would substantially aid the National Library in the establishment of an integrated bibliographic information system.

2. Purchase or rental of a proprietary data base management software package could save the National Library both considerable development time and expense.

3. Software package selection is critically dependent upon the choice of host computer system: most packages will run on IBM 360/370 machines under OS/MVT or OS/VS. There are packages for other machines, but the choice is very much more limited. A first or early step must be to establish the choice of host computer system. Since the host system will presumably be that of the Library and Information Retrieval Centre, this conclusion points to the need for speedy finalization of the plans of the Centre.

4. Even if planning of the Library and Information Retrieval Centre is for some reason delayed, it is to the interest of all future participants to maintain close liaison. Principal aims in this liaison would be:

   (a) establish the present software and hardware commitments of the future participants - without this knowledge effective planning and liaison is difficult

   (b) assist planning of the Centre

   (c) ensure that new hardware, software, program language commitments are as compatible as possible

   (d) exchange information and experience.

   With regard to (c) and (d) above it is pertinent to note that the National Museums have currently reached the benchmark testing stage for certain data base management packages. The possibility exists that the National Library, National Science Library and National Museums could ultimately use at least the same operating system and communications monitor - and possibly also the same data base management system.

Table 3 : Limitations of existing data base management software packages

DBMS	Requirements not satisfied*	Comments
ADABAS	11	Meets most requirements
CAN/OLE	1, 2	Will be studied
DBOMP	1, 2	Only runs on IBM 360/70 under DOS. Literature not studied, additional limitations may be present
Disk Forté	1	Burroughs software product. Literature not studied; additional limitations may be present
DMS-1100	1, 3, 4, 9, 10	Univac software product. Still in development stage. Literature not studied; additional limitations may be present
EDMS	1, 4, 11	XDS software product
GIS	1, 6, 7, 9, 10	IBM software product. Restart and recovery not provided. Must run with IMS to obtain hierarchical and network data structures
IMS	1, 7, 9, 10, 11	IBM software product. Excessive memory requirements. Lacks data independence - change in record format means changes to programs accessing the data. No report generation facility
MARS-III	1	CDC software product. Literature not studied; additional limitations may be present
METABASE	1, 11	IBM 360/370 under OS or VS only. Meets most requirements
SPIRES	1, 3	IBM 360/370 only. Not a proprietary package; software specialists would be required for maintenance. No report generator facility. Literature not studied; additional limitations may be present
System 2000	3, 9, 10	Programs must be precompiled to convert calls Not callable from PL/I
Total	4, 7, 9, 10, 11	Threaded list approach. Utilities to reconstruct data base from log tapes not supplied

* Meaning of numbers given in Table 4

Table 4 : Data base management software requirements

Number	Meaning
1	Operate on a variety of computer makes
2	If IBM, then under OS MVT or OS VS
3	Calls from high level languages
4	Presearch statistics, multi-key search and retrieval
5	On-line updating
6	Variable length records, variable number of variable length fields
7	Empty fields do not require space
8	Growth of data base is open ended
9	New files added to data base without reloading
10	New logical relationships and record structures without reloading
11	Index all types of storage devices (including sequential)

G. RECOMMENDATIONS CONCERNING HARDWARE AND SOFTWARE

1. Planning for the Library and Information Retrieval Centre should be finalized as a matter of urgency. Speedy determination of the host computer system to be employed by the National Library is a matter of the first importance for National Library EDP system planning purposes. Guaranteed access to hardware with suitable operating, communications and data base management software, is a prerequisite for the implementation of an integrated on-line data base serving multiple applications including the CANUC EDP system. Lack of such access would have severe negative implications for National Library EDP system planning and implementation would hinder the integration of National Library EDP applications and would critically weaken* the economic feasibility and practicality of a CANUC EDP system - the centre-piece* of National Library EDP systems to be implemented in the foreseeable future.

2. Liaison between the future participants in the Centre should be strengthened to assist planning, to ensure maximum compatibility and to help the sharing of information and experience. Liaison concerning software development is of particular importance. The various participants in the Centre may commit themselves to different operating systems, communications monitor and data base management software, hindering future operation on a common facility and hindering future integration of files and systems. While a single set of software may not meet all the needs of each participant, liaison should assist agreement on at least common operating system and communication monitor software and on preferred programming languages.

3. The National Library should specify in more detail the data base structure required for its purposes and should carry the study of data base management software packages to the benchmark testing stage. Data base management

---

* Chapter V shows that the CANUC EDP system will have to process a higher volume of records than all the other planned National Library EDP applications put together. It also shows that on-line operations are required for efficient CANUC EDP system performance.

software, such as CAN/OLE, presently used or being developed by other Library and Information Retrieval Centre participants should be studied before final selection of software for benchmark testing.

4. Following testing and selection of data base management software, similar detailed analysis including benchmark tests should be performed in respect of communications monitor packages.

5. Following these stages - and given adequate and economic host computer system availability - the National Library should proceed to implement an integrated data base employing the defined structures and chosen data base management and communications monitor software.

V. THE NATIONAL LIBRARY BIBLIOGRAPHIC DATA BASE: BASIC STRATEGY

The principal factors in the design of the National Library's bibliographic data base are:

1. The requirements of the systems the data base is to support

2. Resources available - in particular, the hardware to be used; software packages available; resources available for software development; development timescale.

Chapter III dealt with timescales; chapter IV dealt with hardware and software; the present chapter outlines the salient requirements of the systems the data base is to support.

A. THE MAIN CURRENT BIBLIOGRAPHIC RECORD INPUTS

Table 5 below shows the main inputs for the years 1974-78, assuming implementation of systems according to the schedule shown in Figure 2.

Table 5 : Projected major current bibliographic record inputs to the National Library data base 1974-78

Subsystem/service[a,b]	1,000 machine readable records: Years commencing 1st April				
	1974	1975	1976	1977	1978
Canadiana[c]	25	34	38	42	46
NL non-Canadiana cataloguing[d]	9	16	23	30	35
NL Acquisitions[e]	-	-	-	-	50
Current US, UK, French, German, Australian MARC records[c]	221	264	301	341	381
ISDS[c]	52	11	16	22	26
CANUC - accessions reports[f]	-	-	-	1,655	1,705
CANUC - location requests[g]	-	-	-	9	13
CANUC - titles[h]	-	-	-	210	216

Notes:

(a) CAN/SDI and CAN/OLE applications are omitted. At present CAN/OLE does not cover the social sciences and humanities. CAN/SDI services related to these subject fields (e.g. ERICTAPES service) do not entail record input to a retrospective base. If CAN/OLE retrospective search services were commenced for records such as US MARC, very serious consideration would have to be given to the means of consolidating CAN/OLE and National Library data base files of these records.

(b) Authority records are not included in the table as it is assumed that they are made up of edited portions of other records shown.

(c) Source: Annex E, p.4.

(d) Source: Appendix 2, p.2.

(e) Subsystem assumed to become operational in 1978. Bibliographic input assumed to equate roughly with NL non-Canadiana cataloguing 1979, plus acquisition of foreign Canadiana.

(f) Source: Annex F, Appendix 1. Serial and non-serial reports have been aggregated. Serials are projected to account for only 20,000 accession reports per year in 1977 and 1978.

(g) Source: Annex F, Appendix 3. Following Table 13 of the Canadian union catalogue location requests survey, only 6% of non-serial location requests are assumed likely to relate to material in the EDP system in its first year of operation, and 12% in its second year. Publication of supplements will lower these figures in practice, and 3% and 6% have been used respectively in the table entries for 1977 and 1978. Following union list publication, serial requests directed to the National Library are assumed to fall to 6,000 per year.

(h) Source: Annex F, Appendix 2.

The important conclusions to be drawn from Table 5 are:

1. The CANUC EDP system

This is the key to any integrated EDP system design for the National Library, since it will have to process a larger volume of records than all the remainder of the National Library's EDP applications put together when it becomes operational in 1977.

2. Data input

The critical feature in the design of the CANUC EDP system is the means by which the input of some 1.7 million accession reports per year is to be achieved. Since an average of 8 locations per title is projected, it follows that after an initial period any accession report will have on average a seven-eighths (88%) chance that its bibliographic data has already been input to the EDP system. The chance will be higher than this if other machine readable records without locations are also on file - for instance, MARC records. In this situation an on-line input system facilitating use of existing machine-readable data will be much more efficient than a batch input system. This argument is crucial to data base design, and is therefore examined in greater detail below.

B. ECONOMICS OF ON-LINE INPUT AND STORAGE IN A HYPOTHETICAL CANUC EDP SYSTEM

Table 6 below analyses the economics of a hypothetical CANUC EDP system employing:

(a) An on-line data base occupying 800 million bytes of disk storage.

(b) 18 terminals to input 1,705,000 accession reports per year - the volume of reports projected in Annex F for the year 1978/79.

The number of terminals is computed, and is not an assumption. Assumptions and computation are explained in greater detail in Appendix 1. Additional background concerning the model system assumed in provided in Annex B. What follows below is a summary of Appendix 1 analysis.

Reports are either keyed in their entirety or input by modifying machine-readable records already in the data base. Alternative assumptions are made concerning the length of records which are keyed and stored.

In "case 1" shown in the table, accession reports of 275 data characters are input and held in the base where they occupy on average 400 characters; the latter figure allows for control characters, directories, data base indexes and an assumed 85% disk utilization. It is much cheaper at $0.12 to modify an existing record than to key an entire record at $1.43. Now, the annual on-line disk storage cost for a 400 character record is only $0.04; in theory, therefore, it would pay to keep a 400 character record over 30 years in on-line storage if it resulted in saving the total keying of another accession report for the same title. The argument is theoretical: in practice, disk rental is not the only cost of keeping records in on-line storage - for example, the data base has to be saved regularly for regeneration purposes and it has to be reorganized. On the other hand, an average of 8 accession reports are expected for each title and this powerfully reinforces the economic benefits of retaining records on-line: on average the keying effort to be saved for each accession report title is 8($1.43 - 0.12) = $10.48.

"Case 2" in the table illustrates what happens to the economics when longer records are keyed and stored: as keying costs are so much greater than storage costs, the case for storing records in order to save keying is even stronger. A record length of 1,000 data characters is chosen as illustrative of the position for National Library cataloguing and Canadiana: these applications employ records of something near this length.

"Case 3" in the table illustrates the position where 1,000 data character records are held in order to save the keying of accession reports of 275 data characters. In theory, it is still worth storing a 1,000 data character record for over a decade in order to save the total keying of a further single 275 data character report for the same title.

Table 6 : Economics[a] of on-line input and storage in a hypothetical CANUC EDP system

Case	Record length in characters		Cost of data entry				Number of decades it is economic to hold records on-line
	No. data characters	Av. length in machine	Modify existing on-line record $	Input new record $	difference $	Annual cost of on-line storage $	
1	275	400	0.12	1.43	1.31	0.04	3
2	1,000	1,235	0.12	5.21	5.09	0.12	4
3	275	1,235	0.12	1.43	1.31	0.12	1

Notes:

[a] The costs shown here include disk storage, terminal rental and labour costs plus overhead. They do not include software, development, maintenance, communication and computer processing costs. For more detailed explanation of assumptions and computation, see Appendix 1. Appendix 1 also shows how the figure of 18 terminals for data input was calculated. For additional background concerning the model system see Annex B and Chapter VI. The figures shown here differ slightly from those in Annex B: the differences are accounted for by use of different throughput assumptions and different keying procedures. The figures here assume Annex F throughput statistics.

Table 6 is illustrative and theoretical, and excludes several categories of cost. The exclusions do not affect the basic conclusions to be drawn provided:

1. Hardware

   Disk storage and computer time are charged on a cost-recovery basis. This is a reasonable assumption in the context of the Library and Information Retrieval Centre.

2. Software and development

   Existing data base management and communications monitor software can be employed - otherwise software and development costs and timescales are unlikely to be acceptable. This is a reasonable assumption in the light of the conclusions of Chapter IV.

3. Communications

   Terminal and communications costs are kept within bounds by ensuring that terminals and tied lines are well utilized. It is assumed that the terminals would initially be in the National Library and National Science Library, and would be connected to the Library and Information Retrieval Centre in the Ottawa area.

With these qualifications it is possible to draw several conclusions from Table 6:

1. On-line input

   An on-line CANUC EDP system operating in a single shift mode would require of the order of 18 terminals for straight input purposes, assuming that all accession reports must be input via terminals.

In fact, a number of projects and institutions such as Ontario Universities' Library Co-operative Project would wish to submit machine-readable accession reports. This would reduce the number of terminals required for on-line input. On the other hand, terminals would also be required for editing and data base maintenance purposes - for instance, authority file maintenance. The important point is that the number of terminals is not excessive, and the system is practical.

2. On-line storage costs

This is no longer the prime constraint on data base strategy, and storage costs are likely to fall (perhaps considerably) over the next few years. At the present time the prime constraints appear to be:

(a) the size of data base that data base management software can handle within acceptable response time limits.

(b) on-line storage available with dedicated access.

While cost is no longer the prime constraint, it is still considerable for a large data base. The difference between the annual costs of storing 275 data character and 1,000 data character records is $(0.12 - 0.04) = $0.08. For a data base of 2 million records this represents a difference in storage cost of $160,000 per year. There is a clear economic advantage at present prices in removing less used records and record data fields from the on-line portion of the base as the base grows.

Against savings achievable by restricting record lengths and record numbers in the on-line base have to be set access delays associated with relegation of data to off-line files, and the processing costs of searching potentially lengthy tapes to retrieve off-line data. The savings to be achieved are lower in the early years of the data base when it is smaller, and in these early years there is likely to be some unused disk storage capacity, since the installation holding the base has to provide for future growth. In addition, a fairly large initial data base is desirable in order to save keying costs and provide support for cataloguing and acquisition processes.

In these circumstances a cost-effective strategy will be to commence on-line data base operation with as many potentially useful records as possible - records containing all potentially useful data fields. Use of the data base should be carefully monitored so that its growth can be controlled in a cost-effective manner. An important aim of monitoring will be to identify lower use records and record data fields. Following identification, these classes of data can be eliminated from the on-line file and placed in off-line storage. This 'purging' of the on-line base file will help to hold down its growth and to keep it within an economic and practical size limit. If a degree of stability in size can be achieved in the first five years of data base operation, there is a very good chance that falling storage costs and improved hardware/software facilities will allow subsequent growth without unacceptable economic and practical penalties.

3. Size of records in the data base

Following the above reasoning MARC records chosen as likely to be useful for cataloguing or accession report input purposes may be held for several years. All the useful data fields may be included in the on-line file: the on-line file can serve many different applications including CANUC accession reporting, CANUC location service, CANUC publication, Canadiana, National Library EDP cataloguing, National Library acquisitions, and catalogue support services for Canadian libraries. A bibliographic item might be entered into the on-line file for one of these applications (say, National Library acquisitions) and subsequently upgraded and used by other applications (say, National Library EDP cataloguing and CANUC).

4. Data base strategy

Summarizing points 2 and 3 above, the National Library should plan to use an integrated on-line data base containing CANUC, National Library acquisitions and in-process files, and selected MARC records, in which the data for each bibliographic item is held only once. All potentially useful records and record data fields should be added to the initial on-line file and monitoring should be used to identify lower use records and data fields. Lower-use records and record data fields should be removed from the on-line file as soon as they have been identified by the monitoring process.

While the data for any title will be held only once in the on-line base, off-line copies of this data will naturally be maintained for regeneration purposes, and some off-line formatted output files will probably also be held in order to save output costs for high volume/ high frequency outputs. It may take some time to integrate all systems, but CANUC EDP system planning should employ an integrated data base concept.

The integrated data base concept is explored more fully in the next section.

C. APPLICATION OF AN INTEGRATED ON-LINE DATA BASE CONCEPT

This concept is illustrated below in Figure 4, which is based on analysis contained in Annex A. The unshaded areas represent off-line files which the National Library may maintain. The shaded area represents the on-line file: the ratio of the shaded area and unshaded area is not to scale and is not significant.

The large totally shaded circle is the on-line "CANUC file" including all union catalogue, union list, and National Library catalogue bibliographic data. The "MARC/ISDS files" circle overlaps with the "CANUC file" circle; overlap represents titles having dual MARC/ISDS and CANUC status. The overlapped titles are all on-line and so are some of the non-overlapped MARC/ISDS titles.

The unshaded portion of the "MARC/ISDS files" circle represents MARC/ISDS bibliographic items held off-line; entries for these items will be included in the on-line indexes - at least for the first years of integrated data base operations. The result of searching the on-line numeric or author/title indexes for one of these titles will be a message to the effect that the title exists in an off-line file. Retrieval from the off-line file will be possible according to some pre-determined schedule. For example, off-line items requested one day might be made available on-line the following day.

Overlapping the two large circles is the small "Files for internal use" circle. Most of these files are on-line, but not all. It will probably be unnecessary to keep a number of non-bibliographic files on-line - for instance, mailing list files. Bibliographic internal files, such as the acquisitions in-process file and the cataloguing in-process file, will be maintained on-line.

Overlapping all of these circles is the central broken line circle representing the central bibliographic authority file. This is an important refinement of the integrated data base concept, although it is possible to use a bibliographic data base without extensive authority files. The Ohio College Library Center data base is an example of a base which as yet does not have a well-developed authority file system.

It is envisaged that the central authority file will be developed to include:

1. Names, references and uniform titles
   Personal, corporate, conference names, uniform titles, and references from alternative forms or spellings of these names and titles to the preferred forms and spellings. Where appropriate, both English and French versions of these would be held.

2. Subjects
   Subject headings in both English and French, possibly linked with associated classification numbers.

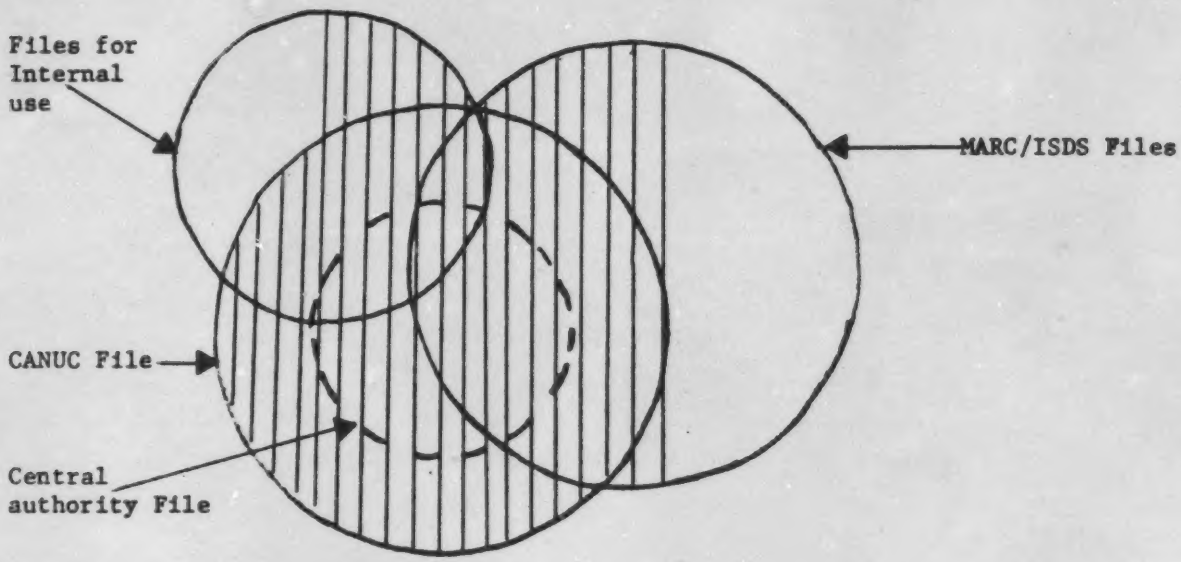
The main purposes of these files are to:

1. Aid consistency in bibliographic recording - particularly cataloguing. A high level of consistency raises search and retrieval efficiency. It is also an important criterion of a high quality catalogue support service.

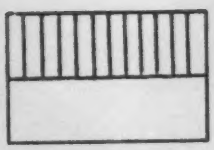
2. Facilitate data base maintenance, correction and search.
   If a given author heading is held once in the authority file rather than in many individual data records, correction is facilitated. The links from the authority file to individual records are also of use for search purposes.

3. Compact the data base.
   Storage space is saved by holding headings once in an authority file rather than replicating them through many individual records.

Figure 4 : Diagram illustrating the National Library integrated data base concept



Legend



- on-line files

- off-line files

D. ORDER OF MAGNITUDE ON-LINE FILE SIZE

It has to be shown that the proposed integrated on-line data base will not become unmanageably large in size. The growth in the size of the largest portions of the data base - CANUC and MARC/ISDS - is illustrated in Table 7 below. It is emphasized that this is only an illustration based on stated assumptions. The basic approach accords with those recommended in Section B of this chapter, but the mode of purging the base to stabilize its size and to save storage costs is somewhat crude, being chosen for simplicity of exposition. The illustrated strategy is as follows:

1. MARC/ISDS records chosen for holding on-line will be ISDS, US MARC, UK MARC, French MARC. Other MARC records will be off-line.
2. MARC/ISDS titles not absorbed into CANUC within three years will be removed from the on-line file. To assist exposition, absorption/purging is shown as occurring in the second year. In fact, the absorption would mostly take place at the end of the third year.

To provide perspective, it is pointed out that the installation on which the national bibliographic data base of France will be stored will have 2,400 million bytes of storage. The table shows that the on-line size will be well within this type of limit for the first ten years of the data base life. The size is also within the range that existing data base management packages can handle. The indicated purging technique is almost certainly not the best choice: monitoring or record use in the first years of operation should be used to establish the optimum record selection technique to be employed.

Practical experience can also assist the choice of off-line storage medium for given categories of record. Some categories of record may be held on magnetic tape and some on off-line disk packs. Use of off-line disk packs is employed successfully at the University of Toronto.

Monitoring will also establish which data fields should be held on-line for lower use items, and which data fields for the same items should be held off-line. For instance, full bibliographic data might eventually be held on-line for records selected as likely to be of use for cataloguing purposes - in general, newer records. At the same time, only short bibliographic data might be held on-line for records selected as being of use solely or mainly for location purposes - in general, older records.

Table 7 : Illustration of the growth in size of the CANUC and MARC/ISDS portions of an on-line data base given an assumed data base strategy

Year Commencing 1st April	Million titles						Million bytes disk storage[e]
	MARC/ISDS			CANUC Input[d]	Total net Annual Input	Data base size	
	Retrospective Input	Current Input[b]	Absorption/ purging[c]				
	A	B	C	D	A+B-C+D		
1977	1.0[a]	0.2	-	0.2	1.4	1.4	900
1978		0.3	0.4	0.2	0.1	1.5	965
1979		0.3	0.5	0.2	-	1.5	965
1980		0.3	0.5	0.2	-	1.5	965
1981		0.4	0.5	0.2	0.1	1.6	1,029
1982		0.4	0.6	0.2	-	1.6	1,029
1983		0.4	0.4	0.3	0.3	1.9	1,221
1984		0.4	0.4	0.3	0.3	2.2	1,415
1985		0.4	0.4	0.3	0.3	2.5	1,608
1986		0.4	0.4	0.3	0.3	2.8	1,800

Notes

(a) Annex E, p.4 shows 1.34 million records; page 9 of the same Annex shows that the number of titles is approximately 77%of this number. 77% of 1.34 = 1.0

(b) Annex E, page 9.

(c) It is assumed that the initial retrospective file will be retained longer than subsequent MARC/ISDS input, since it is not realistic to assume that absorption could run markedly in advance of CANUC input. It would in any case take some time to deal with the 1 million records and time will also be needed to monitor record use. Absorption for the backfile is assumed to take place over the years 1978-82, with final purging during 1982.

(d) Annex F, Appendix 2. In accord with the figures in this appendix, an annual increase of 6000 in the number of titles reported is assumed.

(e) It is assumed that the data base is composed of: 5% non-MARC titles occupying 400 bytes storage each and 95% MARC titles occupying 656 bytes each. This implies an average storage length per record of 643 bytes. The 656 bytes is derived from the average length of US MARC records in distribution format of 636 characters[22], compacted to the extent 78% achieved by Ohio College Library Center, plus the 74 index characters assumed on p.4-2 of Annex B. 85% disk utilization is assumed, as in Annex B.

E. APPROACHES ADOPTED BY SIMILAR PROJECTS ELSEWHERE

The data base strategy outlined above is very similar to the strategies adopted or planned for similar data bases elsewhere. Annex H describes national bibliographic EDP development plans for France and Sweden.

The Bureau pour l'automatisation des bibliothèques, France, plans to add all records of national MARC services to its projected on-line CAPAR (Catalogue partagé) file and to make these available on-line to a number of libraries from 1st January 1975. In 1977 the Bureau will have a new computer installation on the Isle d'Abeau with some 2,400 million bytes of disk storage capacity and 512K core storage.

In the Swedish LIBRIS project

> "All of the university libraries are being connected to a central computer which amongst other things contain a data bank with full bibliographic information on the total stock (books, periodicals, etc.) in the country's academic libraries, together with information on acquisitions, suppliers and references"[23].

In the US, the Ohio College Library Center maintains an on-line file of some 0.8 million titles, nearly half of which are full US MARC records[24] held in a format rather more compact than the US MARC distribution format. The remaining records are those created by OCLC member libraries with as much bibliographic detail as they require for their use.

Apart from broadly common data base approaches these projects are similar in another respect: they are all major long-term enterprises involving considerable expenditures. As noted in Annex H, in 1974 the Bureau pour l'automatisation des bibliothèques will spend its annual operating budget of 3 million Francs ($\simeq$ $0.6 million) plus an equipment allocation of 4.5 million Francs ($\simeq$ $0.9 million): a total expenditure in one year

of some $1.5 million. The long-term nature of this project is indicated by the nature of its planning which spans at least two five-year plans. Ohio College Library Center had a total expenditure of some $860,000 in the year ended 30 June 1973[25]. Although the Center's on-line cataloguing service commenced operation in 1971 it will be several years before the Center's facilities and services cover those already projected. Full development of the Canadian national bibliographic data base and the services supported by this base will be a major enterprise invólving considerable expenditures over the greater part of a decade. Special Canadian requirements, such as those associated with bilingualism and network development in a country of Canada's geographical size, will add to the already complex problems related to the development of a large national bibliographic data base and the services to be provided from this base. The gains of undertaking this type of development may be inferred from the way in which the OCLC system is being replicated in different regions of the USA on a self-supporting basis and from the national plans of France and Sweden, as well as the more direct National Library, CANUC, and other benefits which may be realized in Canada.

VI. FURTHER ASPECTS

A. NETWORK DEVELOPMENT

Chapters IV and V propose the basic central strategy which should guide the planning and implementation of the National Library's bibliographic data base and the systems which this base is to support over the next five years. However, it is important to emphasize that this strategy will need to be developed within a changing national and international bibliographic EDP network context. Such networks are already in existence or taking shape - to mention only four developments:

1. The National Library is already a full participant in the International Serials Data System and the international MARC exchange network.

2. Within 1974 the National Science Library's Canadian On-Line Enquiry (CAN/OLE) pilot project will link some 15 centres across Canada by terminal to the NSL's CAN/OLE data base. The National Library will be one of the participants in this project[26].

3. Within 1974 the National Library will be linked by terminal to the Ontario Universities' Library Co-operative Systems monograph demonstration project, which will have some eight other participating libraries in Ontario and Quebec.

4. Within 1974 the National Library in co-operation with the National Science Library plans to be linked by terminal to the Ohio College Library Center to participate in the US/Canadian co-operative project for the conversion of serials (CONSER).

In this context, conclusions drawn following an examination of French and Swedish network development plans are relevant. The following corclusions are quoted from Annex H:

"1. France and Sweden will achieve their national bibliographic EDP networks much faster and more cheaply than would have been the case if planning had not been attempted.

2. The resulting networks will operate very much more cost-effectively than networks based on uncoordinated development.

3. The case of Sweden shows that planning can be done effectively even though control is not totally centralized. The Swedish mode of planning and network development may have some relevance to countries like Canada, Britain and the United States where control of libraries and information services is relatively decentralized compared with France."

The need for network planning is underlined by the following conclusions drawn in Annex C:

"On-line cataloguing services would be provided more economically by a single national centre than by multiple regional centres, assuming that regional centres replicate national centre hardware and basic facilities ... however, there are other factors besides economics and other EDP services required by Canadian libraries in addition to on-line cataloguing services. A single national centre could not provide all the EDP services required by all libraries in Canada - for instance, such a centre could not have sufficiently fast implementation to provide total selection, acquisition, cataloguing, circulation and IR facilities within the timescales required by Canadian libraries. Also it is not clear that a single national file could or should record the current loan or in-process status of every copy of ever bibliographic item in Canada. As far as comprehensive library automation facilities are concerned, a measure of decentralization is both inevitabl

and a present fact. Beyond catalogue support, therefore, the conclusions of this study merely serve to point to the need for network planning in order to achieve the benefits of co-ordinated development and to minimize duplication of hardware and other facilities. In general it is expensive to replicate items such as mainframe hardware and the software development staff working on the software to be used by the mainframe hardware."

In the light of these conclusions it is recommended that the National Library, in co-operation with other libraries across Canada, should work out a plan for Canadian national bibliographic EDP network development. This plan should be phased and should show for each phase the extent of the network, including:

- details of participating organizations
- location and nature of hardware employed
- location and nature of files accessible to the network
- communications facilities
- services and facilities offered by each participating organization.

This recommendation accords closely with the views of the Canadian Computer/Communications Task Force, exemplified in the following recommendations of the report of the Task Force entitled Branching Out [27]:

1. "Computer/communications (i.e. computer services by remote-access through communications facilities) should be recognized by governments as a key area of industrial and social activity, and steps should be taken towards strengthening ... co-ordination of its development to the benefit of Canadian society."

2. "In the formulation of national computer/communications policy a unified approach throughout Canada should be stressed as a key factor requiring close co-ordination between federal and provincial actions."

3. "In the area of federal responsibilities a Focal Point should be established within the government for co-ordination of the development, formulation and continuing evaluation of national policy in all matters pertaining to the field of computer/communications."

Immediate tasks in the area of bibliographic EDP network planning include:

- the strengthening of organizational arrangements and the provision of additional resources to support joint national/provincial planning concerning library and information service development.

- Canadian National Union Catalogue system planning, for instance the manner in which libraries will report when the new CANUC EDP system is operational. While many libraries may wish to continue reporting exactly as they do now, some will wish to report in machine-readable form. The mode and timescale of this machine-readable reporting needs to be worked out in detail.

- Planning of catalogue support services, both off-line and on-line. Joint national/provincial planning is required since some services will be provided by inter-provincial and intra-provincial projects (for example, Ontario Universities' Library Co-operative System) and some may be provided by foreign or international projects (for example, Ohio College Library Center).

- Active participation in international bibliographic EDP network planning and development, for instance, MARC, ISDS and UNISIST network planning and development.

B. NATIONAL LIBRARY HARDWARE FACILITY REQUIREMENTS

Data base and system strategy recommendations in Chapters IV and V have implications relating to National Library hardware facility requirements. Use of the type of software indicated in Chapter IV implies the following order-of-magnitude core storage requirements:

	Kilobytes
User application programs: allow	250
Data base management software, (e.g. ADABAS 110K))	
Communications monitor software, (e.g. Intercomm 40K))	500
Operating system, (e.g. OS/VS or OS/MVT, allowing 350K))	
	750

Employing the type of data base strategy indicated in Chapter V, disk storage requirements would be of the order of 1,200 million bytes for the first few years of the integrated system. This allows 200 MB for direct access storage of data base management package and operating system software, and assumes that:

(a) The total number of titles in the on-line data base approximates to the total number of projected CANUC, ISDS and French, UK and US MARC titles.

(b) An average of 1,000 data characters per title is sufficient for support of applications implemented in the first few years of integrated system operation.

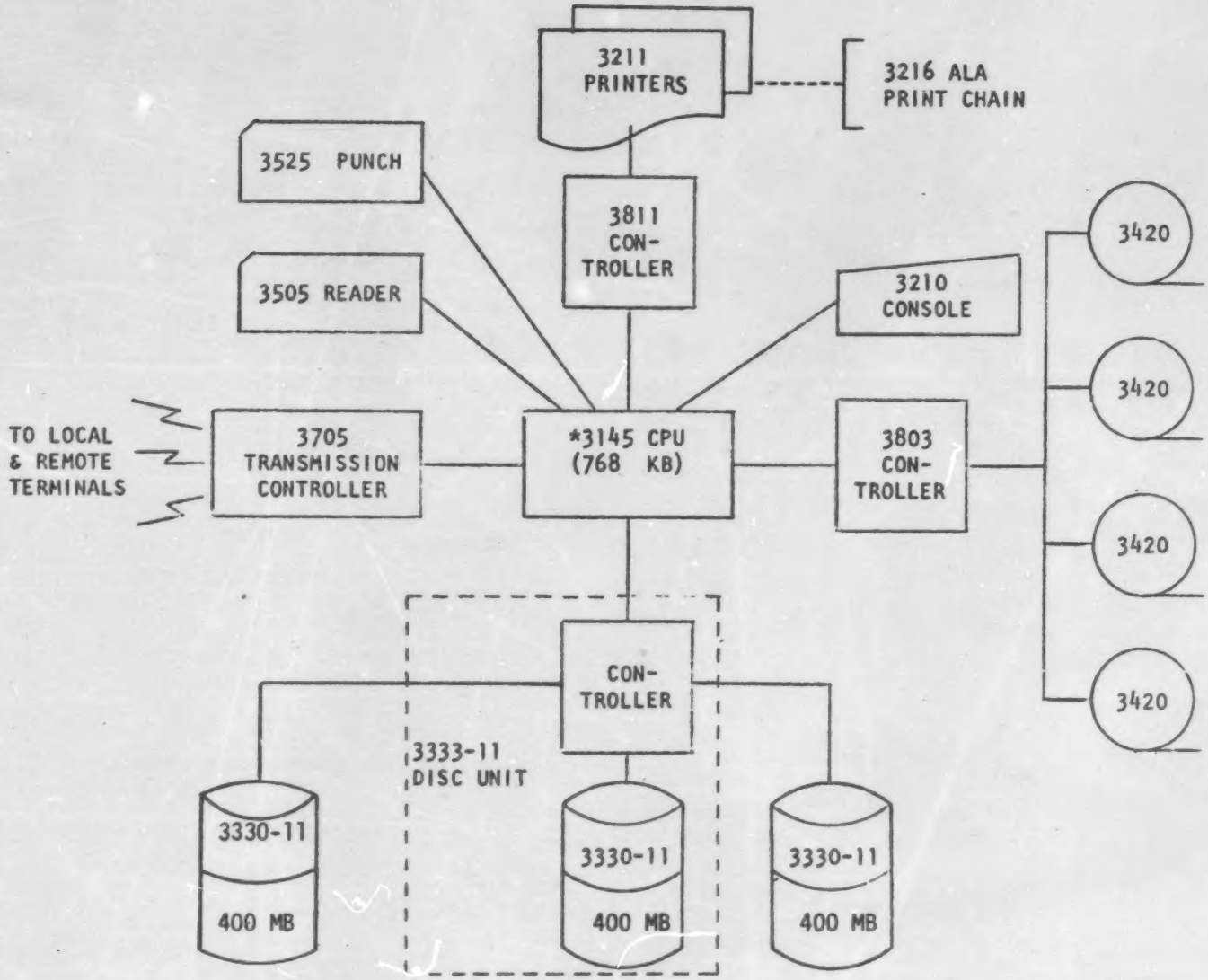
While these assumptions are acceptable for order-of-magnitude projection purposes they will need to be checked when hardware requirements are studied and specified in detail.

Based on these assumptions it is possible to draw out an illustrative computer hardware configuration with the basic facilities needed for adequate support of a National Library on-line integrated bibliographic information system. An illustrative configuration is shown in Table 8 and Figure 5 below; terminals are not shown since the number required depends very much on detailed system design. The configuration employs IBM hardware because this is used by the majority of the organizations which will participate in the Library and Information Retrieval Centre. The Centre would need a larger configuration: Table 8 and Figure 5 indicate the general scope of the facilities the National Library might expect to use within this larger configuration. It is emphasized that Table 8 and Figure 5 should be interpreted in the context of the Library and Information Retrieval Centre: a stand-alone National Library configuration has not been considered - this would require an entire economic feasibility study in itself.

The costings are also illustrative: they are IBM basic monthly rental prices at February 1974 levels. Tendering and contractual terms could make a substantial difference to the final cost of an installation of the type illustrated.

Table 8 : Computer configuration illustrating the basic facilities required for adequate support of a National Library on-line integrated bibliographic information system

Quantity	Description	$/month
1	IBM 370/145 central processer with 768 KB, console and necessary accessories	23,000
1	3333-11 disk unit: controller, plus 400 MB storage	2,530
2	3330-11 double density disk drives with 400 MB each	4,240
1 4	3803 Tape drive controller 3420 Tape drives 800/1600 bpi, 320 KB transfer rate	4,000
1	3705 Transmission control unit	2,000
2	3211 printers with ALA print chains and all necessary accessories	3,300
1	card reader/punch	1,500
	Total	40,570



* = Built-in Controller
for card reader and punch

Figure 5 : Computer configuration illustrating the basic facilities required for adequate support of a National Library on-line integrated bibliographic information system

C. USE AND TREATMENT OF MARC AND ISDS RECORDS

In drawing out a basic National Library bibliographic data base strategy in Chapter V, it was assumed that the National Library would acquire ISDS, US MARC, UK MARC and French MARC records and would add these (initially) to the on-line base. MARC tapes of other countries might be acquired and redistributed, but it was assumed that it was unlikely that there would be economic justification for adding these to the on-line base. The latter assumption was made because languages other than English and French account for a low proportion of use, and a low proportion of material in the base. English and French language material accounts for some 88% location requests; German language material accounts for just over 6% and all other languages account for the remaining 6%[28]. Further, as noted in Annex D, the participants in the Canadian MARC pilot project had no definite requirements for MARC tapes other than those of the US, UK and France. Australian material of prime interest to Canada is assumed to be included in the US MARC tapes.

Addition of (selected) ISDS, US MARC, UK MARC and French MARC records to the on-line base would be useful for several purposes:

1. CANUC accession reports

As indicated in Chapter V these reports require a high-volume input operation. Maintenance of bibliographic data on-line will significantly reduce the otherwise heavy costs of keying reports. ISDS record data can also be used for on-line ISDS/Canada master file purposes.

2. Canadiana/National Library cataloguing

The records will have an on-line catalogue support function for these applications, speeding cataloguing and reducing keying costs.

3. MARC network services

Maintenance of the records on-line will enable the National Library to provide fast response in record selection and provision on behalf of the Canadian MARC network - reducing the need for Canadian libraries to hold large backfiles of MARC/ISDS records. Equally important from the point of view of the network, upgraded records can be supplied. For example, where the National Library has acquired and catalogued a foreign title, request for the foreign MARC record for that title can be met with the record as upgraded and used by the National Library.

These advantages will accrue without an on-line catalogue support network; once such a network service is established, maintenance of records on-line will be mandatory.

A further question relates to the format in which records are held. Annex G indicates that US/Canadian, UK/Canadian and Intermarc/Canadian monographic format machine translations should be acceptable for most practical purposes and that translation costs are unlikely to be excessive. It will, however, take several man-months to specify,and to bring each format translation program to operational status. It is assumed that ISDS and foreign MARC records will be translated to the basic National Library data base format as soon as they are received. Where these records are redistributed, the preferred format will be the Canadian MARC format, as indicated in Annex D.

D. PUBLICATION OF CANUC SUPPLEMENTS

The CANUC EDP system will have the facility of producing supplements to the main union catalogue, where the latter has been published. Publication of supplements will affect the number of location requests directed to the National Library which might be serviced via the EDP system. As shown in Table 5, the number of such requests is too low to affect basic data base

design and strategy. However, it is still of interest to consider the likely cost, mode and pattern of supplement publication. Annex B analyses the costs of publishing supplements in a range of forms: 48x reduction microfiche, 150x reduction ultrafiche, and hard copy.

Costing has to assume definite sequences and publication patterns. Discussions were held with National Library staff to determine the assumptions to be employed, and these discussions led to the selection of a two-section catalogue. The two sections are:

1. Bibliographic data section

   Bibliographic entries in the order of addition to the catalogue, referenced by a running machine-generated number, shown in sample layout in Appendix 3.

2. Index section

   ISBN, LC/Canadiana number and author/title indexes to the master catalogue section. These are also shown in sample layout in Appendix 3.

This arrangement was chosen for several reasons:

* Location searching is speeded since, for any given search, it will in most cases only be necessary to consult one of the indexes.

* Full bibliographic data is provided for resolution of queries relating to requests. This data may also serve a catalogue support function.

* Costs of providing relatively complete bibliographic data are minimized since it is not necessary to cumulate and re-issue the bibliographic data section. It is assumed that additions to this section would be issued at the same intervals as the index section. Alternative index issue frequencies assumed for costing purposes are noted below.

Four cumulation patterns were costed for the indexes:

(a) Issues every 2 months cumulating continuously to a 2 year volume, then starting afresh.

(b) Issues every 2 months cumulating continuously to a 5 year volume, then starting afresh.

(c) Issues every 6 months cumulating continuously to a 2 year volume, then starting afresh.

(d) Issues every 6 months cumulating continuously to a 5 year colume, then starting afresh.

Total subscription costs were calculated for the total catalogue including bibliographic data and index sections. These costs include:

* Computer costs of extracting records from the base, sorting and formatting for printing

* Printing in 300 copies

* Distribution and administration - applied as a percentage to direct costs.

Table 9 shows the total annual subscription to be paid by each of 300 subscribers assuming that catalogue supplements are sold at cost.

Table 9 : Theoretical subscription costs for catalogue supplements under alternative assumptions

Index issue rate	Average yearly per-subscriber cost ($)		
	Microfiche	Ultrafiche	Hard copy
(a) Every 2 months for 2 years	727	879	3400
(b) Every 2 months for 5 years	1616	1606	8949
(c) Every 6 months for 2 years	333	405	1488
(d) Every 6 months for 5 years	629	752	2924

Not visible from the costs is a point pertaining to hard copy. If the issue-every-2-months, volume-every-5-years index-publishing option were chosen, the final issue of the indexes (not including the bibliographic section of the supplements) towards the end of the fourth year would contain 125 thousand pages and would have resulted in a total of two million pages having been progressively superceded over the previous 58 months. Hard copy will, therefore, (in all cases, not just this worst possible one) pose serious doubts concerning the practicality of handling and distribution, and is not recommended.

For the other media, reading equipment would in most cases have to be obtained by subscribers since most libraries do not have 48x or 150x reduction fiche readers. However, the rental cost of readers (under $100 per year for microfiche, and under $300 per year for ultrafiche) is not high enough to affect significantly the choice between hard copy and microform.

Ultrafiche has a distinct handling advantage, since it involves far fewer fiche and is coated with a protective laminate that gives it an expectionally long life. This could be a factor in deciding the media for the main union catalogue, the bibliographic data section of the supplements, and the final cumulation of indexes which will be in constant use for a long period. A counter-argument concerning durability is that since microfiche copies cost only 20¢ each it would be inexpensive to replace damaged or lost microfiche.

The choice between alternative index issue rates involves weighing the advantages of more frequent issues as against the advantages of less frequent issues:

(a) Advantages of more frequent issues.

More frequent issues keep the published union catalogue more up-to-date and allow a higher proportion of location requests to be satisfied via search of the published catalogue. Spreading of interlibrary loan requests can be encouraged by frequent issues, each with a new permutation of the locations associated with each index entry.

(b) Advantages of less frequent issues

Publication costs and subscriptions are lower with less frequent issues, and cumulations can be more comprehensive within a given subscription cost. A less frequent issue rate allows a longer average time for locations to be reported for newly reported titles, facilitating the spreading of the interlibrary loan requests for these titles.

On balance it is suggested that six-monthly issues cumulating to five-year volumes will be a more cost-effective choice than patterns involving issues every two months and/or cumulation to two-year volumes. The principal reason for this suggestion is that a relatively small proportion of requests relate to the newest titles: only about 12% of requests relate to material less than 2 years old[29]. This makes it less important to have very frequent issues, and more important to have longer periods of cumulation.

Reference should be made to Annex B for detailed cost and volume computations and assumptions. The Annex also describes alternative Computer Output Microfilm (COM) processes and equipment and recommends readers for reading 48x reduction microfiche and 150x reduction ultrafiche. The recommendations apply only at the time of writing: in the course of time new readers are likely to become available and there are likely to be changes in the price, delivery, and service positions of existing models. The recommended readers are:

48x reduction microfiche

Bell and Howell, Canada	COM-200	-	$250 (vertical image on screen)
3M Company, Canada	Minicat	-	$239 (horizontal image)

150x reduction ultrafiche

National Cash Register	455-2	-	$725
	455-3	-	$585 (portable reader)

Prices are purchase prices as at February 1974: they are subject to reduction for volume purchase.

Annex B also outlines a possible telephone Voice Answerback enquiry service. This is included as a possibility which might be considered in the longer term.

REFERENCES

1. The EDP master plan; an active plan for the implementation of the electronic data processing policy in the federal government. Ottawa, EDP master plan project, Administrative Policy Branch, Treasury Board Secretariat, December 1972.

2. Ibid., p.27

3. Ibid., p.26

4. EDP systems and projects. Ottawa, Committee on Automation, National Library of Canada, January 9, 1974, 35 pages (typescript).

5. An integrated information system for the National Library of Canada; report of the Systems Development Project. Ottawa, Bureau of Management Consulting Services, Department of Supply and Services, June 1970.

6. Ibid., p.xiii and p.94.

7. National Library News, vol.4, no.6, November-December 1972, pp.6-8

8. National Library News, special issue, January 1974, pp.13-19.

9. Federal Government Library Survey; report of the Services and Systems Team. Ottawa, National Library, 1973, p.37.

10. Ibid., p.288.

11. Reference 8, p.19.

12. Reference 8, p.12.

13. Reference 9, p.60.

14. The final report and recommendations on the establishment of a national machine readable union file of serials for the production of the Canadian Union List of Serials in the social sciences and humanities by Richard Anable. Toronto, York University, 1973.

15. Ibid., p.7.

16. Ibid., p.9.

17. Reference 9, p.151.

18. Reference 8, pp.18-19.

19. "Intercomm", Datapro 70, 1973, p.70E-457-01a. Datapro 70 is published in looseleaf form by Datapro Research Corporation, 1805 Underwood Boulevard, Delran, N.J. 08075. It is available in Canada as Databook 70 through Maclean-Hunter, Toronto.

20. An on-line data base management system for the University of Chicago Library by Jane Jodeit and M. Lee Randall. Chicago, Ill., University of Chicago Library, July 1973, pp.35, 85, 87.

21. "System 2000", Datapro 70, 1972, p.70E-652-01a. See reference 19 for a note on the availability of Datapro 70.

22. Selected statistics based on LC MARC bibliographic records, Systems Office Columbia University Libraries, September 1970 (Technical note no.2).

23. LIBRIS handbook. Part 1. Stockholm, Statskontoret, 1972.

24. OCLC Newsletter. No.66, 28 February 1974, p.1.

25. Ohio College Library Center annual report 1972/73, p.16.

26. National Science Library Bulletin, No.12, January 1974, p.5.

27. Branching out; report of the Canadian Computer/Communications Task Force. Ottawa, Department of Communications, May 1972, Vol.I, p.203.

28. Canadian national union catalogue location requests survey, Ottawa, Research and Planning Branch, National Library, August 1972, p.34.

29. Reference 28, p.33.

ECONOMICS OF ON-LINE INPUT AND STORAGE IN A HYPOTHETICAL CANUC EDP SYSTEM

The purpose of this appendix is to show the assumptions and computation used to give the figures presented in Table 6 : Economics of on-line input and storage in a hypothetical CANUC EDP system. The hypothetical CANUC EDP system is that detailed in Annex B with accession report volume amended to accord with Annex F projections.

1. Disk storage occupied by records and indexes

As detailed in Annex B, page 4-2, each record is assumed to require indexes of 74 characters in length. This figure is not assumed to increase for longer records since all essential access points are covered in short records and the longer records simply have more data - not more index access points.

In addition, disk utilization is assumed to be only 85%, i.e. records carry a 15% data storage overhead. A further overhead is assumed in the next paragraph. This assumption is pessimistic: Ohio College Library Center compact US MARC records to 78% of their communication format length without loss of data. They are able to reconstitute the records in full if necessary.

These assumptions result in the following average record storage lengths for data records of 275 and 1,000 characters.

data record characters	computation	average record storage length
275	(275 data characters + 74 index characters) $\frac{115}{100}$	400 characters
1,000	(1,000 data characters + 74 index characters) $\frac{115}{100}$	1235 characters

2. Cost of disk storage

A data base size of 800 million bytes (MB) is assumed, occupied by data (600 MB) and software (200 MB). The software storage overhead is included in total data base storage costs, i.e. the cost for 800 MB storage is charged against the data which occupies only 600 MB. The software storage overhead increases data storage unit costs more in a small base than in a large one. The assumption of a base which is relatively small compared with the 2400 MB disk storage projected for the installation to hold France's national bibliographic data base is made because conservative rather than optimistic costings are preferred.

1 x 3333-11 (Controller with 400 MB)	$2,530/month
1 x 3330-11 (Dual density; 400 MB)	$2,120/month
	$4,650/month

If 600 MB of data is stored for $4,650/month = 55,800/year, then:

100 bytes of data are stored for	$\$0.93 \times 10^{-3}$ per year
400 bytes of data are stored for	$\$3.72 \times 10^{-3}$ per year
1235 bytes of data are stored for	$\$11.48 \times 10^{-3}$ per year

3. Terminal operation

A terminal operator costs $15,000 per year including overhead. His terminal costs $2,400 per year. He works 220 days per year, keying 4,000 keystrokes per hour. He therefore achieves 4.4 million key depressions per year. Assuming single-shift operation, the 4.4 million key depressions cost $15,000 + $2,400 = $17,400 per year. A lower cost per million key depressions would almost certainly be achieved by use of more than one shift, if this could be arranged. Use of more than one shift is not assumed, in case union or other difficulties prevented it.

4. Assumptions for terminal computation

1,705,000 accession reports are projected for the year 1978/79. Since it

is projected that the average title will be reported eight times, 7 out of 8(88%) reports will find matches. Assuming that a selection of MARC records with zero locations is also held, an even higher proportion will find matches. This proportion in percentage terms is assumed to be 95%.

350 key depressions are required to input full details of records not on file: 275 characters plus 65 characters for shifts, field identifiers, etc. and a further 10 characters for additional data added because it occurs on the accession report and appears useful.

Records already on file are assumed to require 20 key depressions to find the existing machine readable report, 4 to add a new location and a further 6 for additional data or editing, i.e. a total of 30 key depressions. Additional data is assumed to be added where early accession reports contain less data than subsequent accession reports.

5. Terminals required for hypothetical CANUC EDP system

Assuming that all accessions reports are entered in a single shift operation by terminals, the above assumptions lead to the following computation:

(a) Input of records with bibliographic data already on file:

$$\frac{95}{100} \times \frac{1{,}705{,}000 \text{ accession reports} \times 30 \text{ key depressions}}{4.4 \text{ million key depressions per terminal/year}} = 11.04 \text{ terminals}$$

(b) Input of records without bibliographic data already on file:

$$\frac{5}{100} \times \frac{1{,}705{,}000 \text{ accession reports} \times 350 \text{ key depressions}}{4.4 \text{ million key depressions per terminal/year}} = 6.78 \text{ terminals}$$

Rounded Total 18.0 terminals

6. Unit costs of inputting an accession report by modifying an existing machine readable record

Following the above computation, in 1978/79 there will be 95% of 1,705,000 = 1,620,000 additional locations to existing titles which will occupy 11 terminals and operators @ $17,400.

$$\text{Unit cost per additional location} = \frac{\$17400 \times 11}{1{,}620{,}000} = \$0.1181$$

7. Unit cost of keying total accession report

In 1978/79 there will be 5% of 1,705,000 = 85,000 accession reports being input from scratch and occupying 7 terminals

$$\text{Unit cost per totally keyed accession report} = \frac{\$17400 \times 7}{85000} = \$1.433$$

The above unit cost assumes records of 275 basic data characters. A record of 1000 basic data characters will have proportionately higher unit costs.

$$\text{Unit cost of input of 1000 character input} = \$1.433 \times \frac{1000}{275} = \$5.211$$

8. Unit cost of all accession report input

1,705,000 accession reports occupy 18 terminals and operators.

$$\text{Overall unit cost} = \frac{\$17400 \times 18}{1{,}705{,}000} = \$0.184$$

PROJECTED NATIONAL LIBRARY NON-CANADIANA MACHINE READABLE CATALOGUING 1974-78

The purpose of the projection presented in this appendix is to help indicate:

1. cataloguing use of the data base in respect of non-Canadiana National Library material

2. likely storage requirements in the data base.

ASSUMPTIONS

1. NL non-Canadiana cataloguing output

The following figures are drawn from the Forecast of output volumes for the 1975/76 program forecast for the Descriptive Cataloguing and Canadiana Editorial Divisions:

	1972/3	1977/8	% increase
(a) Full and partial cataloguing: Canadiana and General	59,609	119,110	100
(b) Entries to Canadiana	25,561	42,000	64
(a) - (b), i.e. Full and partial cataloguing: General	34,048	77,100	126

The 126% increase in "General" cataloguing over a five year period implies a compound annual rate of increase of 17.7%. The figures relate to volumes, copies and entries. Figures for titles provide a more meaningful basis on which to project the growth of the National Library's machine readable data base. For this reason, the projection below relates to figures for titles catalogued, drawn from page 46 of the Annual Report of the National Librarian 1972-3. An annual rate of increase of 13% compound has been assumed, allowing for the possibility that funding and other problems may cause actual performance to lag behind budget figures. The result is shown below:

	1972/3	1977/8	1978/9
Cataloguing for the National Library Catalogues Full and partial cataloguing Non-Canadiana titles	16,592*	30,000	34,500

2. NL non-Canadiana machine readable cataloguing

This commenced in January 1974 and is assumed to cover all NL non-Canadiana cataloguing output by 1 April 1977. The growth of the EDP system coverage is assumed to be linear.

PROJECTIONS

Year commencing 1st April	Thousand records
1973	2
1974	9
1975	16
1976	23
1977	30
1978	35
Total	115

*This figure was obtained by subtracting Canadiana figures for monographs, printed music, audio-visual materials and microforms from total National Library figures for the same materials, since the figures for National Library titles in the Annual Report of the National Librarian 1972-73 include Canadiana cataloguing. Estimates by the National Library staff were obtained in order to determine the number of periodicals catalogued solely for the general collection.

TENTATIVE HYPOTHETICAL EXAMPLE OF POSSIBLE BIBLIOGRAPHIC DATA SECTION

001 Valéry, Paul, 1871-1945
... Autres rhumbs. [Paris] Gallimard, 1934. 266p.
ISBN 000-321-147-X
ISBN 000-321-149-1

34-2894

002 Vangi, Giuliano, 1931-
Vangi Magistrat der Stadt Darmstadt, Kunstverein Darmstadt Italienishes Generalkonsultat Frankfurt, Kunsthalle Darmstadt, 27 März bis 9 Mai 1971 [Katalog Redaktion und Gestaltung Bernd Krimmel, 1971?] 61p.
(Meister der italienischen Modern, 10)
I. Krimmel, Bernd, ed. II. Kunsthalle Darmstadt.
ISBN 12-678-3421-9

72-34435

003 Vacuum technology at low temperatures
S. Alexander Stern, editor, [by] Randall F. Barron and others. New York, American Institute of Chemical Engineers, 1972. 65p. (AICRE symposium series no.125 v.68)
I. Stern, S Alexander, ed. II. Barron, Randall. III. American Institute of Chemical Engineers. AICRE symposium series, no.125.
ISSN 0432-8765
ISBN 01-4783-784

72-8961

004 Varona y Pera, Enrique José. 1849-1930
Textos escogidos. Ensayo de interpretacion, acctaciones, y seleccion por Raimundo Lazo [1st ed.] México, Editorial por rua, 1968. XXV, 162p.
(Sepan cuantos, no.25)
I. Lazo, Raimundo, ed.

005 Smith, Julia Floyd, 1914-
Slavery and plantation growth in Antebellum, Florida, 1821-1860. Gainsville, University of Florida Press [1973].
IX, 249.
ISBN 0-8130-0823-7

70-15065

006 Smaridge, Norah
You know better than that. Drawings by Susan Perl. Nashville, Abingdon Press [1973] [82]p.
I. Perl, Susan, illus.
ISBN 0-687-46744-6

72-7033 MARC

Note: This is a sample that does not necessarily conform to layout standards.

SAMPLE ISBN INDEX

ISBN	CANUC No.	Locations
0-13-936237-1	024	QMM
0-13-936245-2	024	QMM
0-395-12665-7	017	NSY OKIT OD QMQ SSU NSKS OK NFSG NBS BVI
0-517-50020-5	015	OOS OTB NSHDIP QMFRAN QSHERSH
0-517-50021-3	015	OOS OTB NSHDIP QMFRAN QSHERSH
0-687-46744-6	006	OOP QMJ OTMCL
0-695-40047-9	009	AC ACLS BVA MW NBS NFSG NSHRL OB QMJ SR YWR
0-695-80047-7	009	AC ACLS BVA MW NBS NFSC NSHRL OB QMJ SR YWR
0-8027-6123-2	010	AC YWR NSHRL BVA NBS MW OB SR QMJ NFSG
0-5027-6124-0	010	AC YWR NSHRL BVA NBS MW OB SR QMJ NFSG
0-8130-0323-7	005	OONL QMM OTV

SAMPLE LC/CANADIANA INDEX

LC/Canadiana No.	CANUC No.	Locations
34-28944	001	OOU OOGE OPOC QMAY OTULS QMMBZ QQLAM QSHERSS BVAUW AE NFSQ NSHPD OOMR
68-57116	012	OOA OONL QMM QMU QMU BVIV
70-150656	005	OONL QMM OTU
70-153957	017	NSY OKIT OD QMQ SSU NSKS OK NFSG NBS BVI
71-131065	016	QML OOP OLK
72-8	024	QMM
72-2499	009	AC ACLS BVA MW NBS NFSG NSHRL OB QMJ SR YWR
72-7033	006	OOP QMJ OTMCL
72-81380	010	AC YWR NSHRL BVA NBS MW OB SR QMJ NFSG
72-84318	015	OOS OTB NSHDIP QMFRAN QSHERSH
72-89619	003	OOUM YWR NSDB

SAMPLE "AUTHOR/TITLE" INDEX

AUTHOR & TITLE, TITLE, ADDED ENTRIES, SERIES	CANUC No.	LOCATIONS
AIChE Symposium Series, No. 125, V8.....	003	OOUM YWR NSDB
American Institute of Chemical Engineers. AIChE...	003	OOUM YWR NSDB
Anglistica & Americana, No.71	013	OONL
Asimov, Isaac. Comets and meteors.....	009	AC ACLS BVA MW NBS NFSG NSHRL OB QMJ SR
Asimov, Isaac. How did we find out about elec....	010	AC YWR NSHRL BVA NBS MW OB SR QMJ NFSG
Asimov, Isaac. The stars in their courses.....	016	QML OOP OLK
Asimov, Isaac. Treasury of humor.....	017	NSY OKIT OD QMQ SSU NSKS OK NFSG NBS BVI
Autres rhumbs.....	001	OOU OOGE OPOC QMAY OTULS QMMBZ QQLAM QSHERS
Barron, Randall F. Vacuum technology at low temp..	003	OOUM YWR NSDB
Bartlett, Percy W. International economic re-....	014	OPET
Bullokar, John. An English Expositor.....	013	OONL
Comets and Meteors.....	009	AC ACLS BVA MW NBS NFSG NSHRL OB QMJ SR
Diesel Engineers and Users Association. Pub.342...	025	OKA OLUL OCKA
Dow, George Francis. Slave ships and slaving.....	012	OOA OONL QMM QMU OTU BVIV
Economic Appeasement.....	023	OOCC
English Expositor, An.....	013	OONL
Etruscans in an ancient world the.....	018	MBC